

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:38:59 ; Search time 48.4 Seconds
(without alignments)
1559,432 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGLVYLGALALAGLGFPPAD.....PSPTPGSTLTPPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

- Database :
- 1: SP-archaea:*
 - 2: SP-bacteria:*
 - 3: SP-fungi:*
 - 4: SP-human:*
 - 5: SP-invertebrate:*
 - 6: SP-mammal:*
 - 7: SP-mhc:*
 - 8: SP-organelle:*
 - 9: SP-phage:*
 - 10: SP-plant:*
 - 11: SP-rodent:*
 - 12: SP-virus:*
 - 13: SP-vertebrate:*
 - 14: SP-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2654	91.0	468	4	Q9UC32
2	1840	63.1	577	11	Q35370
3	1396.5	47.9	461	11	P97883
4	583	20.0	652	4	Q9NPY3
5	580.5	19.9	757	4	Q9HCU0
6	543	18.6	652	4	O00274
7	530.5	18.2	644	11	O89103
8	528.5	18.1	643	11	O9ET61
9	528.5	18.1	643	11	O9J126
10	433	14.8	1664	5	Q9VVO2
11	414	14.2	1574	11	O88281
12	385	13.2	1394	5	Q9V589
13	375	12.9	708	13	P87363
14	374	12.8	2906	11	Q9WUH9
15	366	12.6	3857	11	O88840
16	360	12.3	528	11	O9CXDB
17	360	12.3	2872	11	Q9WUH8
18	359	12.3	741	4	O9NSZ1
19	358.5	12.3	1174	11	Q99K58

20	353.5	12.1	1242	4	Q9NS15	Q9NS15 homo sapien
21	353.5	12.1	1382	4	Q9H7K2	Q9H7K2 homo sapien
22	351.5	12.1	1511	4	O75412	O75412 homo sapien
23	351.5	12.1	1587	4	O00508	O00508 homo sapien
24	341	11.7	576	4	O9Y3V7	O9Y3V7 homo sapien
25	338	11.6	1062	11	O60789	O60789 mus musculu
26	333.5	11.4	961	11	O9E0C6	O9E0C6 mus musculu
27	333	11.4	999	4	Q9N036	Q9N036 homo sapien
28	331	11.4	589	5	O9T2S1	O9T2S1 caenorhabd
29	330.5	11.3	956	11	Q99K64	Q99K64 mus musculu
30	330	11.3	2189	5	O9E055	O9E055 elmeria ten
31	329.5	11.3	1664	11	O35442	O35442 mus musculu
32	328.5	11.3	1713	11	O88349	O88349 mus musculu
33	328.5	11.3	1764	11	O35806	O35806 raltus norv
34	327.5	11.2	1095	11	O60784	O60784 mus musculu
35	323	11.1	1253	11	O61810	O61810 mus musculu
36	320.5	11.0	937	5	O9BLJ1	O9BLJ1 clona intes
37	320.5	11.0	2321	4	O9Y6L8	O9Y6L8 homo sapien
38	320.5	11.0	2321	4	Q9UM47	Q9UM47 homo sapien
39	319.5	11.0	1833	11	O08999	O08999 mus musculu
40	315.5	10.8	2281	4	Q90PL3	Q90PL3 homo sapien
41	315	10.8	997	11	O9JUS0	O9JUS0 mus musculu
42	314.5	10.8	3507	5	O23587	O23587 caenorhabd
43	313	10.7	495	4	O9HRO5	O9HRO5 homo sapien
44	313	10.7	798	5	O18026	O18026 caenorhabd
45	312.5	10.7	1821	4	O14767	O14767 homo sapien

ALIGNMENTS

RESULT 1

Q9UC32 PRELIMINARY: PRT: 468 AA.

Q9UC32: 01-MAY-2000 (TREMURel. 13, Created)

01-MAY-2000 (TREMURel. 13, Last sequence update)

01-JUN-2001 (TREMURel. 17, Last annotation update)

THROMBOMODULIN.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

[1]

SEQUENCE.

MEDLINE=9293792; PubMed=8390446;

Yanamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;

"Urinary thrombomodulin, its isolation and characterization.";

J. Biochem. 113:433-440(1993).

HSSP: P07204; IZAO.

InterPro: IPR00152; Asx_hydroxyl.

InterPro: IPR000561; EGF-like.

InterPro: IPR001881; EGF_Ca.

InterPro: IPR001304; lectin_C.

InterPro: IPR001491; Thrombomodulin.

Pfam: PF00008; EGF_5.

DR	Pfam: PF00059; lectin_C_1.
DR	PRINTS: PR00907; THROMBOMODULN.
DR	SMART: SM0034; CLECT; 1.
DR	SMART: SM00181; EGF_6.
DR	PROSITE: PS00010; ASX_HYDROXYL; 1.
DR	PROSITE: PS00041; CTYPE_LCTIN_2; 1.
DR	PROSITE: PS01186; EGF_2; 2.
DR	PROSITE: PS01187; EGF_CA; 1.
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ	SEQUENCE 468 AA; 49444 MW; 4BFE898EF86A40 CRC64;

Query Match 91.0%; Score 2654; DB 4; Length 468;

Best Local Similarity 99.6%; Pred. No. 8 4e-204;

Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q9 19 APEAPGSGSQCVEHDFALYGPATFLMASQICDGLRGLMTYRSSVAADYISLLINGD 78

```

Db 1 APAEPQSGSCVCHDFCALYPGPATFLNLSQICDGLRGLMTVRSSVADVLSLLNGD 60
QY 79 GGVGRRLMIGLQLPKCGDPRKRLGRLGFQWNTGDNNTSYSKMARLDINGAPLCPICV 138
Db 61 GGVGRRLMIGLQLPKCGDPRKRLGRLGFQWNTGDNNTSYSKMARLDINGAPLCPICV 120
QY 139 AVSAEAATVSEPTWEBOQCEVKADGFLCEHFPAICRPLAVPBGAAAASVTYCTPRA 198
Db 121 AVSAEAATVSEPTWEBOQCEVKADGFLCEHFPAICRPLAVPBGAAAASVTYCTPRA 180
QY 199 ARGADFOALPYGSSAAVAPLGLQIMCTAPPGAVOGHMAREAPGAMPSCVENGCCEHACNA 258
Db 181 ARGADFOALPYGSSAAVAPLGLQIMCTAPPGAVOGHMAREAPGAMPSCVENGCCEHACNA 240
QY 259 ITPGARQCCPAGALQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKETGYRLAA 318
Db 241 ITPGARQCCPAGALQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKETGYRLAA 300
QY 319 DQHCEDVDDCILPSPCPQRCVNTQGGFECCHCYPNVDLVDGECVPPVDFCFRANCETOC 378
Db 301 DQHCEDVDDCILPSPCPQRCVNTQGGFECCHCYPNVDLVDGECVPPVDFCFRANCETOC 360
QY 379 QPLNQTSTLCYCAAGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEGYIIDDGFT 438
Db 361 QPLNQTSTLCYCAAGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEGYIIDDGFT 420
QY 439 CTDIDECENGGFCGVCCHNLPGTFECICGPDLSALVRIHIGTDCDSCGYVD 486
Db 421 CTDIDECENGGFCGVCCHNLPGTFECICGPDLSALVRIHIGTDCDSCGYVD 468
RESULT 2
O35370 PRELIMINARY; PRT; 577 AA.
AC 035370:
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF022743; AAB80760.1; -.
DR HSP: AF022742; AAB80923.1; -.
DR HSP: P07204; 1FGD.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_5.
DR PRINTS: PR00059; lectin_c.1.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_Ca; 2.
DR SMART: SM00001; EGF_Like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_Ca; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

```

```

SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BFB18555F CRC64;
Query Match 63.1%; Score 1840; DB 11; Length 577;
Best Local Similarity 63.8%; Pred. No. 8e-159;
Matches 332; Conservative 49; Mismatches 133; Indels 6; Gaps 5;
QY 1 MLGLVLVIGALALADLPAPAPAFPGSGSCVCHDFCALYPGPATFLNLSQICDGLRGLM 60
Db 1 MLGLVLVIGALALADLPAPAPAFPGSGSCVCHDFCALYPGPATFLNLSQICDGLRGLM 60
QY 61 TVRSSVADVLSLLNDGCGRRRLMIGLQLPKCGDPRKRLGRLGFQWNTGDNNTSY 120
Db 61 TVRSSVADVLSLLVDSMSDRP-WIGLQLPQGGDPVHLGRLGFPQWNTGDNNTSY 118
QY 121 RMARLDINGAPLCPICVAAATVSEPTWEBOQCEVKADGFLCEHFPAICRPLAV 180
Db 119 RMARPNQSPPLCPICPLTVTATEAPRPAEERPCENETKGFCEFFFAAFCRPLV 178
QY 181 EP-GAAAASVTYCTPFAARGADFOALPYGSSAAVAPLGLQIMCTAPPGAVOGHMARE 239
Db 179 NTRPDEGAHISSTYNTPLVYSGADFOALPYGSSAAVAPLGLQIMCTAPPGAVOGHMARE 238
QY 240 PGAMDCSVENGCCEHACNAIPGARQCCPAGALQADGSRCTASATQSCNDLCEHFCVNP 299
Db 239 TGAWNCSEVENGCCEHACNANRANGRCVCPGSGDLQADGSRCAKAPVQLCNELCQHFVNN 298
QY 300 PDQGSYSCKMETGYRLAAOHRCEDVDCCILPSPCPQRCVNTQGGFECCHCYPNVDLVD 359
Db 299 SDVSGTSSCKMETGYRLAAGHRCEDVDCCILPSPCPQRCVNTQGGFECCHCYPNVDLVD 358
QY 360 GECEVPYDFCFRANCETOCPLNQTSTLCYCAAGFAPIPHEPHRCMFCNQACPADCDP 419
Db 359 GECEVQLDPCFRSKCEYQCPVNSTHNCIACGAFAPKDDPDRCMFENETSCPADCP 418
QY 420 NTQASCECEGYIIDDGFTCTDIDECENGGFCGVCCHNLPGTFECICGPDLSALVRIHIGT 479
Db 419 NSPFCQCPREFIIDECSICTDIDECSQGECLINECRNLPGTFECICGPDLAQAQISKD 478
QY 480 CDSGKV--DGDGSGSGEPSPS-PTPGSTLTPPAVGLVHSG 516
Db 479 CDPPLVLEDSGSGGHPSSNPTVYSSITVPFARRMHSG 518
RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
AC P97883:
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE THROMBOMODULIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang J., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: U90121; ABA49723.1; -.
DR HSP: P07204; 1FGD.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_5.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00179; EGF_Ca; 2.
DR SMART: SM00001; EGF_Like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS01186; EGF_2; 3.

```

DR PROSITE: PS01187; EGF_CA: 2
KM Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
SQ SEQUENCE 461 AA; 49499 MW; 45971BCE84688E67 CRC64;

Query Match 47.9%; Score 1396.5; DB 11; Length 461;
Best Local Similarity 62.4%; Pred. No. 1.5e-103;
Matches 244; Conservative 37; Mismatches 105; Indels 5; Gaps 4;

QY 131 PLGGLPLCVASAAEATVPSPETWEEOCEVKADGLCEHFRPNTC-RPLAVER-GAAAAA 188
DB 12 PLGGLPLCVTVSTTEAPGEPAAEKEPCENETGFLCEFTFAAFCSRPLVNTTRDPEGAH 71
QY 189 VSTTYTPFPAAGADFOALPVGSSAAVAPLGLQLMCTAPGAVOGHAREAPGAMDCSVE 248
DB 72 ISSTYTPFAGVSADPOTLPFGSSATVAVAPGLVLCALGCTSEGHWTRVETGAMNCSVE 131
QY 249 NGGCEHACNAIPGAPRCOCAPAGALADGRSCATASATQSCNDLCEHFCVNPDPQGSYSC 308
DB 132 NGGCEHACNMSANGPRCVCPSGGDLADGRSCAKPVGQLCNELCQHFCVNNNDVPQGSYSC 191
QY 309 MCTETGRLAADQRCEDVDICILEPSPRCQVNTQGGFECHCYPNVDLYDGEVPEVDP 368
DB 192 MCTETGYQLAADGHRCEBDVDCCKQGNPCPOLCSNTEGGFRCRCYDGYELVDGECVEQLDP 251
QY 369 CEFANEEYOCPLNONTSYLCVCAEGFAPRHEPHRCOMFCNOTACPADCDPNTQASCECP 428
DB 252 CFSKCEYQCQPVNASTHYNTICAEGRFAPKLDDPRCEMFCNETISCPADCPNPSFCQCP 311
QY 429 EGYILDGFICTIDIDECENGSGVCHNLPGTFECICGPDALVRHIGTDCDSGKY--D 486
DB 312 EGYILDGSGICTIDIDECSGQEBCLTNECRNLPGSYECICGPDALAGQISKDCDIPVLED 371
QY 487 GGDGSGEPPPS-PTPEGSTLTPPAVGLVHSG 516
DB 372 SEDGSGEHPSSNPTVSVTPPSARPMSHG 402

RESULT 4
Q9NPY3 PRELIMINARY; PRT; 652 AA.
ID 09NPY3;
AC 09NPY3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE D3737E23.1 (EGF-LIKE DOMAINS CONTAINING C10/MBL/SPA RECEPTOR (C10R)).
GN D3737E23.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RA Sehra H.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL118508; CAC00597.1;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF Ca.
DR InterPro: IPR001304; Ilectin_C.
DR Pfam; PF00008; EGF_5.
DR SMART; SM00034; CLDCT; 1.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00001; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Receptor; Repeat.
SQ SEQUENCE 652 AA; 68559 MW; EECADFAC55FCAC2 CRC64;

Query Match 20.0%; Score 583; DB 4; Length 652;
Best Local Similarity 31.8%; Pred. No. 1.5e-38;
Matches 181; Conservative 55; Mismatches 210; Indels 124; Gaps 32;

QY 2 LGVLVGLALALA--GLGFPAPAEPOGSGOCEVDEHDCFPALYPAPTFPLNASQICDGLRGL 59
DB 5 MGLILLILLILLITOGGAGTADTE----AYVGVGACTAHAGKLSAAEAQHCHQNGNGL 60
QY 60 MTVRSSVAAD----VISLLNGDGVGR--RLMIGLQLPPG-CGDKRRLPLRGFOWT 112
DB 61 ATVASKKEAQHVQGVNLQLLRREALATPARMSKFWNLGQREKGLDPSL--PLKGFVWG 118
QY 113 GDNNTYSRRMARLDLNGAPLCPGLCA--VSAAEATVSE-PIMEEOC-----EVAD 163
DB 119 GGEDTPYSNMHKELRNSC--ISKRCVSLIDLISQPLPSRLPKWSEGGSGSPGSGNIE 176
QY 164 GLCEFHFRPACRPLAV-EPGAAAASITYGTPPAAGADFOALPVGSSAAVAPLGLQL 222
DB 177 GFVCKFSFKGKCRPLALGPG-----OVTYTPPQTSSSLAEAVPFASANVA----- 224
QY 223 MC-TAPGAVOGHW---AREAPGAMD-----CSVENGCEHAC-NAIPGAP 263
DB 225 -CGEGDKDETQSHYFLCKEKAPDVFDWSSGRLCVSPKYGCFNNGGCHQDCEFGDGSF 283
QY 264 RCQCPAGALADGRSCATASATQSCNDLCE--HFCVNPDPQGSYSCMETGYRLADQH 321
DB 284 LCGCRPGRLLDLDLVTC-ASRNPCSSPSRGATCVLP-HQKNYTCRCRPGYOLDSSQL 341
QY 322 RCEVDYDCILPSPRCQVNTQGGFECHCYPNVDLYDGEVPEVDPGFRANCEYQCP 381
DB 342 DCVVDVDC--QDSPAQCCVMTWPGFRCWCWGY-----EPGCP-----GEGACQDV 386
QY 382 NOTSYLCVCAEGFAPRHEPHRCQMFNOCFACPADCDPNTQAS--CEGEGYTL--DDGF 437
DB 387 DE-----CALGRSP-----CAQGC-TNTDSFRCSEEGVILAGEDET 423
QY 438 ICTDIDEC--ENGSGVCHNLPGTFECICGPDALVRHIGTDCDSGKY-----D 486
DB 424 QCQDVDECVGPGRLDCLCFNTQGSFHGCLPGLVLAIPN-GVSCMGVSLGPPSGPBD 482

QY 487 GGDGSGE-----PPSPTPGSTRPFA 509
DB 483 EEDKGEKGSIVPRAATASPTRGPGTPKA 512

RESULT 5
Q9HCU0 PRELIMINARY; PRT; 757 AA.
ID 09HCU0
AC 09HCU0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TUMOR ENDOTHELIAL MARKER 1 PRECURSOR (ENDOSTALIN PROTEIN).
GN TBM1 OR ENDOSTALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20407466; PubMed=10947988;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lai A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT *Genes expressed in human tumor endothelium.*;
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT *Molecular cloning and characterization of Endostalin, a C-type
lectin-like cell surface receptor of tumor Endothelium.*;
RT

RL J. Biol. Chem. 276:7408-7414(2001).
 DR EMBL: AF279142; AAG00867.1; -.
 DR EMBL: AJ295846; CAC34381.1; -.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 3.
 DR SMART: SM00179; EGF_Ca; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Signal.
 FT CHAIN 1 17 POTENTIAL.
 FT SIGNAL 18 757 TOMOR ENDOTHELIAL MARKER 1.
 FT SEQUENCE 757 AA; 80859 MW; C96363BA1FD8FFA0 CRC64;

Query Match 19.9%; Score 580.5; DB 4; Length 757;
 Best Local Similarity 33.4%; Pred. No. 2.7e-38;
 Matches 176; Conservative 44; Mismatches 192; Indels 115; Gaps 24;

QY 1 MGVLYVGLALALAGLG-FPAPAPROPGSGCYEHDCFAIYPGATFLNASQICDGLRHL 59
 DB 2 LRLILLAMAAAGPTLGODPWAAPR---AACGSPSCYALFPRRRRTLEAMRACRELGGDL 58
 QY 60 MYRVSSVAADVISLLINGDGVGRRRLMIGLO.PPGCGDPKRLRGROWVTDNNTSY 119
 DB 59 ATRTRTEERORVSLVG--AGPASKRLMIGLOQANOCLOLR--PLRGTTWTTGDDDTAF 114
 QY 120 SRMARLDNGAPLGPICVAVSAEATVPSEPIWBOGCEVKAAGFLCEFHPRATCRPLA 179
 DB 115 TNNAQ-PASGSGPCAPORCALEAS-----GEHMLEGSCSTLAVDYLQGFEGACPALQ 168
 QY 180 VEEGAAAAAVSITYGTPFAKAGDFQALPVGSSAAV---APLGLQIMCTARPAVAGHNA 236
 DB 169 DEAGQAGPAV---YTPRFILVSTEFEMLPFGSYAAVQCAAGRASLLCYKQREGVGW-WS 224
 QY 237 REAP---GAWDCSVENGCEHAC-NAIPGARPCQAGALADAGSCVATASATQ--CNLD 291
 DB 225 RASPLIGCT-GCSPDNGCCHCEVEVDGHSVRCIEGFLADAGSCDEPCQACQCEQ 283
 QY 292 CEHFCAVNPDPGSGSCMETGTGRLAADQ-HRCEVDVDCITLSEPCPCRCVNTGSGFECH 350
 DB 284 CE-----PGPGQYSGCHRIAGFPRAEDPHRCVDTDECOI-AGVQOQMCVNYVVGFECEY 336
 QY 351 CYNYVDLVDECEVPVDFPCFRANCEYQCCQPLNFTSLCYAEGEFAPIPHPHRCQMFQNG 410
 DB 337 CSBGHELE-----ADGIS----- 349
 QY 411 TAGPADCP---NTQASCCEPBGYIILDGFICTIDE--CENGFGCSGVCHNLPGT-- 461
 DB 350 -----CSPAGANGAQAQSLDGL-ELLDDGEDEDEDEAKAFANGMT-----EMGILM 397
 QY 462 FEICIGDSDALVRHIGTDCDSKVDGSDSGEPSPPTPGSTLTPP 508
 DB 398 MEPTOPDFALAYRPSFEPD-----REPQIYPEPTWPPP 432

RESULT 6
 ID 000274 PRELIMINARY; PRT; 652 AA.
 AC 000274;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE C1QR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NBL_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97199258; PubMed=9047234;
 RX Neimincene R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
 RT "cDNA cloning and primary structure analysis of C1QR(P), the human
 C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro."
 RL Immunity 6:119-129(1997).
 DR EMBL: U94533; AAB53110.1; -.
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF; 5.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00179; EGF_Ca; 3.
 DR SMART: SM00001; EGF-like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_Ca; 3.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW SEQUENCE 652 AA; 68576 MW; B7EAB5F5714A775 CRC64;

Query Match 19.9%; Score 580; DB 4; Length 652;
 Best Local Similarity 31.6%; Pred. No. 2.5e-38;
 Matches 180; Conservative 56; Mismatches 210; Indels 124; Gaps 32;

QY 2 LCVLYVGLALALA-GLPAPAPROPGSGCYEHDCFAIYGPATFLNASQICGLRHL 59
 DB 5 MCLILLILLILLPQAGGAGATE---AVCVGTCYAAHSGKLSAEMAHCHNONGNL 60
 QY 60 MYRVSSVAAD---VISLLINGDGVGRR--RLMIGLOLPFG-GGDPKRLPLRGPOVNT 112
 DB 61 AIVKKEEQAQVRLADQLRREALTARMKSFNIGLOREKGLDPSL-PLGFGFWG 118
 QY 113 GDNNTSYSMAALDINGAPLGPICVA--VSAEATVPSE-PWEEQOC-----EYKAD 163
 DB 119 GSEDPIYSWMHKELEHNSC--ISKRCVSLILLDSOPLLNRLPKWSEGGSPGSGSNIE 176
 QY 164 GFICEFHPRATCRPLAV-EPGAAAASITYGTPFAKAGDFQALPVGSSAAVAPLGLQL 222
 DB 177 GFVCKFSEKMGCRPLALGGP-----OYTYTPPQTSSLEAVPFASANVA----- 224
 QY 223 MC-TAPGAVQGHM--AREAPGAMD-----CSVENGCEHAC-NAIPGAR 263
 DB 225 -CGEDDKDETQGHYFLCKEKAADVDFDMGSSGPLCVSPRYGCNPNNGGCHQDCFGGDSNF 283
 QY 264 RCQCPAGAAALQADGRSCTASATQSCNDICE-HFCVPPNDPGSTSCCTGTGRLAADQH 321
 DB 284 LCGRRPGEFLDLDLYTC-ASRNPCCSSSPCRGATCVLGP-HGKNYTCRCPOGYLDSSQL 341
 QY 322 RCEVDVDCITLSEPCPCRCVNTGSGFECHCYPNVDLVGECVEPVDFPCFRANCEYQCCPL 381
 DB 342 DCVVDVDEC-QDSPAQCECVNTPGGFRCCEWGY-----EFGGP-----GEGACDV 386
 QY 382 NOTSYLCVCAEFAPILPHPHRCQMFQNOTACPADCDPNTQAS--CECPBGYIIL--DDGF 437
 DB 387 DE-----CALGRSP-----CAQGC-TNTDGSFRCSEEGYVLAGEGGT 423
 QY 438 ICTDIDEC--ENGFGCSGVCHNLPGTFPICGPGSALVRHIGTDCSGKV-----D 486
 DB 424 QCQDVDECVGPGPLCDLSLCEFTQGSFHCGLPGVLA PN-GVSCTMPVSLGPPSPGPPD 482
 QY 487 GGDGSGE-----PPSPPTPGSTLTPPA 509
 DB 483 EBDKKEGSGTVPRAATASPTRGPECTPA 512

RESULT 7
 OX 89103

ID	AC	PRELIMINARY;	PRT;	644 AA.
DT	01-NOV-1998	(TEMBLrel. 08, Created)		
DT	01-NOV-1998	(TEMBLrel. 08, Last sequence update)		
DT	01-JUN-2001	(TEMBLrel. 17, Last annotation update)		
DE	C1Q/MBL/SPA RECEPTOR C1QRp.			
DE	ly68 OR C1QRp OR A44.			
OS	Mus musculus (mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVEV;			
RA	Tenner A.J., Kim T.S.;			
RT	"Identification of the mouse genomic DNA for C1QRp."			
RT	"Molecular characterization of A44, an early marker of hematopoietic development."			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RA	Peterenko O., Lemischka I.R.;			
RT	"Molecular characterization of A44, an early marker of hematopoietic development."			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SV;			
RA	MEDLINE=99359842; Pubmed=10430665;			
RT	Noteworthy P.J., Taylor P.R., Walport M.J., Botto M.;			
RT	"Cloning of the mouse homolog of the 126-kDa human C1q/MBL/SP-A receptor, C1QRp."			
RL	Mamm. Genome 10:789-793(1999).			
DR	EMBL: AF074856; AAC63274.1; -			
DR	EMBL: AF081789; AAC62649.1; -			
DR	EMBL: AF099939; AAD47906.1; -			
DR	EMBL: AF099938; AAD47906.1; JOINED.			
DR	HSSP: P07204; 1EGT.			
DR	MGD: MGI:106664; ly68.			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001881; EGF_Ca.			
DR	InterPro: IPR001304; Lectin_C.			
DR	Pfam: PF00008; EGF_5.			
DR	Pfam: PF00059; Lectin_C; 1.			
DR	SMART: SM00034; CLECT; 1.			
DR	SMART: SM00179; EGF_CA; 3.			
DR	SMART: SM00001; EGF_like; 2.			
DR	PROSITE: PS00010; ASX_HYDROXYL; 3.			
DR	PROSITE: PS50041; C_TYPE_LECTIN_2; 1.			
DR	PROSITE: PS01186; EGF_2; 3.			
DR	PROSITE: PS01187; EGF_CA; 3.			
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Receptor; Repeat.			
KW	SEQUENCE 644 AA; 69354 MW; EB4351648BF835A CRC64;			
SO	SEQUENCE			
Query Match	18.6%	Score 543;	DB 11;	Length 644;
Best Local Similarity	30.08;	Pred. No. 2.2e-35;		
Matches 167;	Conservative 55;	Mismatches 211;	Indels 124;	Gaps
5	LVLGALTAGAGCGFPAAPAPRPGGSGCYEHDFALYPPAPATFLNLSQICDGLRGHTLVRS	64		
9	LILGLGLGPMWG--AAADSQ--AVVCESTACTYTAHMGKLSAAEAQHRCNENNGMLATVKS	64		
65	SVAA-----DVLSILNGCGYGR--RLMTGLQRPFGCGDPKRLGRLRGQVYTGDNNTS	118		
65	EEEAHHVOALTLQLLKTARPLEAKMGKRWIGLOREKGGCTYHDL--PKRGFSWGGGSDTA	123		
119	YRRMRRLDNGCAPLGLCPICVAVSAAEATVPE--PIWEBOCEV-----KADGFLCEHF	171		
124	YENMWKAKSSGCTIFRCVSLILDSLTIRPHSHLRKMHESPCGTPEARAGNSIEGTLCENF	163		
172	PATCPRLAV--EPGAAAAAVSTITGTPRPAKARDQALPVGSSAAVAPRLQD-----LM	223		

Db	184	KGMCRLPLAGSGG-----RVYTTTPTFOATTSSLEAVPPASVANYA-CGDPAKSETHYFL	236
Qy	224	CT--APRGAVQGMHAREAP-----GAMDCSYVNGGCEIAC-NAIRGAPRPGCCAPGAALADAGS	277
Db	237	CNKRTPGIF--HWSSSGGLVSPKPFSGSFNNKGCQDDDEGGDGSFRCGCPGFRLLDDI	294
Qy	278	RCSATASATQSCNDLCEHFVCVNPDPQ-----SYSCMCEGYRLADQHRCEVDY	327
Db	295	VTCAS-----RNCSSNRPCTGGGCMCHSVPLSENITCKPCSGYGLDSSYVHCNCDID	344
Qy	328	DCILEPSPCPCRCVNTQGGFECCHCYPNYDLVDS--ECVEPYDPCFRANCEYQCQPLNOT	384
Db	345	EC--QDSPCADCCVNTLGSFHCBCWVQY-PSGPKKEACEEDVDECAANSP-----	392
Qy	385	SYLVCACAEFARIPIEPIHRCMFCNCPACADDPDTQAS--CECPREYIL--DDGEICT	440
Db	393	-----CAQCT-----NTDGSFSCSKREYIYSGEDSTQCE	423
Qy	441	DIDECEN--GGFCSCGVCHNLPTGETECIGDPSALVHNIGTDCDSG-----KYD	486
Db	424	DIDECSDANGNCDLSCINTGTGSPRCGCPGMELAPN--GVFCSRGTVSELPAPRPQKED	482
Qy	487	GGDSSGSGPEPPSPYSGS	503
Db	483	NDRKESTYMPTEMPSS	499
RESULT	8		
Q9ET61			
ID	Q9ET61	PRELIMINARY:	PRT: 643 AA.
AC	Q9ET61		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	C1Q/MBL/SPA RECEPTOR C1QRP PRECURSOR.		
CN	C1QRP.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PIV;		
RA	Lovik G., Vaage J.T., Dissen E., Sepier C., Ryan J.C., Rolstad B.;		
RT	"Characterization and molecular cloning of rat C1qrp, a receptor on		
RT	macrophages, natural killer cells and dendritic cells."		
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF136537; AAG0157.1.		
DR	Interpro: IPR000152; Asx_hydroxyl.		
DR	Interpro: IPR000561; EGF-like.		
DR	Interpro: IPR001881; EGF-like.		
DR	Interpro: IPR001304; lectin_c.		
DR	Pfam: PF00008; EGF. 5.		
DR	SMART: SM00034; CLECT. 1.		
DR	SMART: SM00181; EGF. 5.		
DR	SMART: SM00179; EGF_CA. 5.		
DR	SMART: SM00001; EGF-like. 2.		
DR	PROSITE: PS00010; ASX_HIDROXYL; UNKNOWN_1.		
DR	PROSITE: PS50041; C_Type_Lectin_2; 1.		
DR	KW Signal; Receptor.		
FT	SIGNAL 1 23 POTENTIAL.		
SO	SEQUENCE 643 AA; 68781 MW; 9AE6C93AD943DB6 CRC64;		
	Query Match	18.2%; Score 530.5; DB 11; Length 643;	
	Best Local Similarity	29.7%; Pred. No. 2..2e-34;	
	Matches 166; Conservative 56; Mismatches 209; Indels 127; Gaps		
Qy	4	VIVVIGALALAGLGFAPAPAPPGSGSQCVEHDCALYVGPATFLMASQICDGLRGHTMTVR	63
Db	8	LILVLGLDGLQIMAGCAADSE---AVVCGSTACTYAHNGKLSAAEAQHRCHNENGGNLTATVK	63
Qy	64	SSVA-----DIVSILLN-----GDGGVGRRRRLMTGLDLPCCGDPKRLGLRGFGQVWTGDN	115

```

Db 64 SEERHVOEALADLKTAPSETYIG--KEWIGLQREKGCYHDL-PMKGFSWVGGE 120
QY 116 NTSYSMARLUNG--APLCGPLCAVSAAEATVPSE--PIWEOQCEV-----KADGFL 166
Db 121 DTTYSNMWKAKSSSISKRCSVLIDSLKPH--PSHLPKWHESPCCGPDAFGNSIGFL 178
QY 167 CEHFEPATCRPLAV--EFGAAAANVITYGTPPAARGADFOALPVGSSAAVAPLGLQ---- 221
Db 179 CKFNFGKCSPLALGGPG-----QLYTTPPQATTSLSLKAVPASVAVNY--CGDEAESK 231
QY 222 ---LMCTAPPAVVOGHMAREAP---GAMDCSVENGCEHAC--NAIPGAPRCOCAGAL 273
Db 232 TNYVLCCKETTAGV--FHWSSGFLCVSPKFGCSFNNGCQDCCFEGGDSFRCGRPFRL 290
QY 274 QADGRCTASATQSCNDLCEHFVCVNPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTCTAS-----RNPCCSNPCTGGCMCHSVPLSENYYCHCPRGYQDLSQVHC 340
QY 324 EDVDDCLLEPSPQRCVNTQGFECCHCYPNYDLVDG--ECVEPVPDPCFRANCERYOQPL 381
Db 341 VDIDEC--EDSPQDQECINTPGSFHCEGWYQSSGSKKEACEDVDEC----- 386
QY 382 NOTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQAS--CECPBGYIL--DGF 437
Db 387 --TAAYSPCAOG-----CT-----NTDGSFYCSCKEGYIMSGDKST 420
QY 438 ICTDIDECENGFCGVCNHLPGTFECICGPDALVHRIGTDCDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRCGCPAGFELAPN--GVSCIRGSMFSELPAHPQKE 478
QY 486 DGDGSGSGEPSPSPGSS 503
Db 479 DKGDGKESTVPLTEMPGS 496

RESULT 9
Q9JITZ6 PRELIMINARY; PRT; 643 AA.
AC Q9JIT6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE C10RP.
GN C10RP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN RP SEQUENCE FROM N.A.
RC STRAIN=MISTER; TISSUE=LUNG;
RA Dean Y.D., McGreal E.P., Akatsu H., Morgan B.P., Gasque P.;
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: A1160978; AAF80402.1.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00041; C_TYPR_LCTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 3.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 643 AA; 68780 MW; F823D742E02A4931 CRC64;

```

Query Match

18.1%; Score 528.5; DB 11; Length 643;

```

Best Local Similarity 29.7%; Pred. No. 3,2e-34;
Matches 166; Conservative 55; Mismatches 210; Indels 127; Gaps 28;

QY 4 VLVLALALAGLGPAPAPQPGSGQCEHDCALYATPFLNAAQIDGLHMLTVR 63
Db 8 LLLGLGLQWAGAAASE-----AVYEGTACTAHMGKLSAEADHRCNENGNINATVK 63
QY 64 SSVA-----DIVSLLN-----GDGVRRLRLIGLDPRECQDPKRLGRLRGFWYTGDN 115
Db 64 SEERHVOEALADLKTAPSETYIG--KEWIGLQREKGCYHDL-PMKGFSWVGGE 120
QY 116 NTSYSMARLUNG--APLCGPLCAVSAAEATVPSE--PIWEOQCEV-----KADGFL 166
Db 121 DTTYSNMWKAKSSSISKRCSVLIDSLKPH--PSHLPKWHESPCCGPDAFGNSIGFL 178
QY 167 CEHFEPATCRPLAV--EFGAAAANVITYGTPPAARGADFOALPVGSSAAVAPLGLQ---- 221
Db 179 CKFNFGKCSPLALGGPG-----QLYTTPPQATTSLSLKAVPASVAVNY--CGDEAESK 231
QY 222 ---LMCTAPPAVVOGHMAREAP---GAMDCSVENGCEHAC--NAIPGAPRCOCAGAL 273
Db 232 TNYVLCCKETTAGV--FHWSSGFLCVSPKFGCSFNNGCQDCCFEGGDSFRCGRPFRL 290
QY 274 QADGRCTASATQSCNDLCEHFVCVNPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTCTAS-----RNPCCSNPCTGGCMCHSVPLSENYYCHCPRGYQDLSQVHC 340
QY 324 EDVDDCLLEPSPQRCVNTQGFECCHCYPNYDLVDG--ECVEPVPDPCFRANCERYOQPL 381
Db 341 VDIDEC--EDSPQDQECINTPGSFHCEGWYQSSGSKKEACEDVDEC----- 386
QY 382 NOTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQAS--CECPBGYIL--DGF 437
Db 387 --TAAYSPCAOG-----CT-----NTDGSFYCSCKEGYIMSGDKST 420
QY 438 ICTDIDECENGFCGVCNHLPGTFECICGPDALVHRIGTDCDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRCGCPAGFELAPN--GVSCIRGSMFSELPAHPQKE 478
QY 486 DGDGSGSGEPSPSPGSS 503
Db 479 DKGDGKESTVPLTEMPGS 496

RESULT 10
Q9TVQ2 PRELIMINARY; PRT; 1664 AA.
AC Q9TVQ2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Y6AG10A.7 PROTEIN.
GN Y6AG10A.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN RP SEQUENCE FROM N.A.
RC Mortimore B.J.;
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 262:2012-2018(1998).
[3]
RP SEQUENCE FROM N.A.
RA Ainscough R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
EMBL: A117206; CAB60454.1; -

```


RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agrawal A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Balle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benson P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Meperson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshneff A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheffler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA GIBBS R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003558; AAF50538.1; -.
DR HSSP: P07204; 1ADX.
DR FlyBase: FBgn0035798; CG7526.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF_Ca.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_12.
DR Pfam: PF00084; sush1; 2.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00179; EGF_Ca; 9.
DR SMART: SM00001; EGF_Like; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; 8.
DR PROSITE: PS01186; EGF_2; 9.
DR PROSITE: PS01187; EGF_Ca; 10.
KM Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1394 1394
SQ SEQUENCE 1394 AA; 152269 MW; CD29380E3162P68A CRC64;

Query Match 13.2%; Score 385; DB 5; Length 1394;
Best Local Similarity 32.1%; Pred. No. 1.9e-22;
Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;

QY 224 CTAPGAVOG-----HWAPAGAW-----DCSVENGCEHCNATIPG 261
DB 563 CLCPGATGALGIDNHTIVSLNSFTIDTSSETPSAHNTCLDIDECSLANGNSHFQNPBG 622
QY 262 APRCCPAGALQADGRCSTASATQSC--NDLCENHCVPNDPQDGSYSCMETGYRLAA 318
DB 623 GFGCACPGYALSDMRCC--QDIDECIDNSNGQCSQLCL--NQGCGFACACETGFELTP 677
QY 319 DQHCEDVDCTILPSPQRCVNTQGGFEGHCFTPNYDLVGE--CVPVPDPC--FRAN 373
DB 678 DFGCADIDECSDQYGNCSIDICINLGHACACERGYELAKDKLSLD--VECCAGILSGG 736
QY 374 CEYCCOPLNPF--SYLCVABEGFAPRPHRCQMFQNCQFACPA--DCDPRNQASC----- 425

DB 737 CSNHC--INKAGTECCGPIG--ILNDGK-----SCSPALVGGCPPTGORSADCCAP 785
QY 426 -ECPGTYLDDGFTCTDIDEC--ENGFGCSGVCHNLGTEPCITGCP-----DSALVRHI 476
DB 786 IECNPGTYLSDDKCVDIDECQKONG--CSHRCNTEGSEKSCPPGEYLDSDQKTCODI 844
QY 477 GTDCDSGK 484
DB 845 -DECDQDK 851
RESULT 13
ID P87363 PRELIMINARY; PRT; 708 AA.
AC P87363;
DT 01-MAY-1997 (TREMURel. 03, Created)
DT 01-MAY-1997 (TREMURel. 03, Last sequence update)
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)
DE FTBRILLIN-1 (FRAGMENT).
GN FBNI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou G., Price C., Godfrey M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U88872; ABA46531.1; -.
DR HSSP: P07204; 1RGD.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF_Like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF; 14.
DR Pfam: PF00084; TB; 2.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00179; EGF_Ca; 14.
DR PROSITE: PS00010; ASX_HYDROXYL; 13.
DR PROSITE: PS01186; EGF_2; 10.
DR PROSITE: PS01187; EGF_Ca; 13.
KM Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
FT NON_TER 708 708
SQ SEQUENCE 708 AA; 76163 MW; C247271CIDF73361 CRC64;

Query Match 12.9%; Score 375; DB 13; Length 708;
Best Local Similarity 31.8%; Pred. No. 6e-22;
Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

QY 244 DCSVENGCSEHCNATIPGAPRCQCPAGALQADGRCSTASATQSCNDLCENHCVPNP--- 300
DB 223 ECSIWNGCSEHCNATIPGAPRCQCPAGALQADGRCSTASATQSCNDLCENHCVPNP--- 300
QY 301 -----DQPGSYSCMCEGYNRLAADQHRCEVDVDCILPSPC--PQRCVNTQGGFEGHCYR 353
DB 273 DGGQCTNIPGEGRCICLYGFMASEDKTCVDNBECDLHPNCLSGTCENTGSGFICHQDM 332
QY 354 NYDLVGE--CVPVPDPC--FRANCEYOCPLN--QTSYLCVABEGFA-----PIP 398
DB 333 GYSGKKGTTGCTD--INEEIGAHNCDRAVCTNIPGSKSCSSSWINGIKCTDLDSCS 391
QY 399 HEPRHCQMFQNTACPADCPNTQAS--CEPEGYILDDGFTCTDIDEC-----CENG- 448
DB 392 NGTHKCSPH-----ADC-KNTWMSYRCLCEGY--TGDGFTCTDLDCESENTLNCENGQ 442
QY 449 -----GF-----CS-----GVCNHLPGTFECLCGDSALV 473
DB 443 CLNAPGRCBCDMGFLPSLDGKACEDIDECSLPNICVYGCCHNLPGLFGECEVGEYELD 502
QY 474 RHIG-----TDCDSGK--VDGSDSGSGEPPPTPSTLTPRAVGLV 513

```

Db          503  RSGNGCTDVNECADPPTTCIGTCVNTAGSYTCECPDP-----FELNPTRVGCV 550
RESULT      14
ID           09MUH9      PRELIMINARY;      PRT;      2906 AA.
AC           09MUH9;
DT           01-NOV-1999 (TReMBLrel. 12, Created)
DT           01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT           01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE           FIBRILLIN-2.
OS           Rattus norvegicus (Rat).
OC           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX           NCBI_TaxID=10116;
RN           [1]
RP           SEQUENCE FROM N.A.
RA           Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N.,
RA           Wallner E.I., Kanwar Y.S.;
RT           "Cloning of rat fibrillin-2 cDNA and its role in branching
RT           morphogenesis of rat embryonic lung."
RL           Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR           EMBL; AF135060; AAD34439.1; -.
DR           HSSP; P35555; 1EMN.
DR           InterPro; IPR002086; Aldehyde_dehydr.
DR           InterPro; IPR000152; Asx_hydroxyl.
DR           InterPro; IPR000561; EGF-like.
DR           InterPro; IPR001881; EGF_CA.
DR           InterPro; IPR001438; EGF_II.
DR           InterPro; IPR002212; TB.
DR           Pfam; PF00008; TGF_46.
DR           Pfam; PF00683; TB; 9.
DR           PRINTS; PRO0010; EGFBL00D.
DR           SMART; SM00179; EGF_CA; 42.
DR           SMART; SM00001; EGF_Like; 4.
DR           PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR           PROSITE; PS00010; ASX_HYDROXYL; 43.
DR           PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR           PROSITE; PS01186; EGF_2; 36.
DR           PROSITE; PS01187; EGF_CA; 43.
KW           Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat
KW           SEQUENCE      2906 AA; 313572 MW; 9EE64E727044EF58 CRC64;

```

Query Match	12.8%;	Score 374;	DB 11;	Length 2906;
Best Local Similarity	-32.9%;	Pred. No. 2.9e-21;		
Matches 102; Conservative	24;	Mismatches 92;	Indels 92;	Gaps 18

OY	244	DCSYENGGCHACNAIIPGARCCCPAGAALOADRGCTTASATQSCMDLCEHFCVNPDP--	301
Dd	1237	ECMIINMGCGTCOTNSGSGTECCSSGYALMPDGRSCA-----DIDE--CENNPPIC	1286
OY	302	-----QPSSYSCMETGYRLAADOHRCEDVDCLLEPSPCP-ORCVNTQGFEFCHCYR	353
Dd	1287	DGGGCTNIPEHYKGLCIDGFMAIMDKTKCIDVNECDLNIMCFEGCEWTKSFICHOL	1344
OY	354	NYDLVDGE-CVEPVDP-C-FRANCETOCPLN-QTSYLCAEGBFA-----PIR	398
Dd	1347	GYSVKKGATGCTD-VDECEIGAINCOMHAASCLNVPSFSCREGVWGNGIKCIDLDEBA	1405
OY	399	HEPHRCMFCNOTACPADCDPTQAS--CECEPGYLLDDGFTCTDIDE-----CENG-	448
Dd	1406	NGTHOCSI-----NAQC-VMTPGSVRCASCSEGF-TGDGFTCSIDVDECAENINICEHQ	1455
OY	449	-----GF-----CS-----GVCHNLPRFTECTCGPDALV	473
Dd	1457	CLANPYARCECEMGFTPASDRSCODIDECSEFONICVFGTCONNLPRGMHCICDDGYAD	1516
OY	474	RHIG--TDCC	481
Dd	1517	RTGGHCTDID	1526

RESULT	15			
088840				
ID	088840	PRELIMINARY;	PRT;	3857 AA.
AC	088840;			
DT	01-NOV-1998 (TrEMBLrel. 08, Created)			
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	MUTANT FIBRILLIN-1.			
GN	FN1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=B10.D2;			
RC	MEDLINE=98069008; PubMed=9405934;			
RA	Bona C.A., Mural C., Casares S., Kasturi K., Nishimura H., Honjo T.,			
RA	Matsuda F.,			
RT	"Structure of the mutant fibrillin-1 gene in the tight skin (TSK)			
RT	mouse.";			
RL	DNA Res. 4:267-271(1997).			
DR	EMBL: AF007248; AAC62317.1; .			
DR	HSSP: P35555; IAPJ			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR002557; Chitin_biding.			
DR	InterPro: IPR000561; EGF_1like.			
DR	InterPro: IPR001881; EGF_Ca.			
DR	InterPro: IPR002212; TB.			
DR	InterPro: IPR000822; znf-C2H2.			
DR	Pfam: PF00008; EGF_64.			
DR	Pfam: PF00683; TB; 12.			
DR	SMART: SM00494; ChEBD2; 2.			
DR	SMART: SM00179; EGF_CA; 20.			
DR	SMART: SM00001; EGF_1like; 4.			
DR	PROSITE: PS00010; Asx_HYDROXYL; 61.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE: PS01186; EGF_2; 50.			
DR	PROSITE: PS01187; EGF_CA; 61.			
DR	PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.			
DR	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat			
DR	SEQUENCE 3857 AA; 418503 MW; 5BC0618BC527ED04 CRC64;			

Query Match	12.6%;	Score 366;	DB 11;	Length 3857;
Best Local Similarity	30.3%;	Pred. NO. 1.7e-20;		
Matches 118;	Conservative 28;	Mismatches 132;	Indels 112;	Gaps 21;

OY	212	SAAVAPIGL-----QUMCTAPGAVAGHARAPAMCMCSVENGCCEACNAITGARRC	265
Db	2148	SANLCSPHGRCVNIIGKIQACMGNYHPHTDRLEFCVIDIDCISIMNGSECFETCNISGDISEC	2207
OY	266	QCPCGALOADGRSCTASATASCNDICEHCPCVPN-----DOPGSQSCMCTGYRL	316
Db	2208	SCQGFFLMPDQNSCT-----DDIE--CEDNPICDGGOCTNPGETRCLCYDGFMA	2257
OY	317	AADHRCEEDYDDCILERPSPC-PORCVNTQSGFECHCYPNWDLVDVE--CVEPEVC--FR	371
Db	2258	SEDKKTCVDVNIEDLNINICLSTCENTGYSFICHCDMKYSGKRKGTTGTD-INCELTGA	2316
OY	372	ANCEYOCPINOT-SYLVCABEGFA-----PIPEHPRCOMFCNOTACPADCNP	420
Db	2317	HNCDRHAIVCNTTAYGSEFCSSPGWMIDGKITCTDLDECSSGTMMCSQH-----ADC-KN	2366
OY	421	TQAS--CECPBGITLLDGEICTDIIBC-EN-----GGF-----	450
Db	2369	TMGSEIQLCKIDXY-TGDFGFTCTDLDECSENTNLNCGNGOCLNAPGRYBCDMGFVPSADG	2427
OY	451	-----GS-----GVCHNLPGRFEELICGSDSLVNHIG-----TDCCSGK	484
Db	2428	KACEDIDECIPNVICFGTCIHNPJRGFERCECEIGEILDSGGNCITDVNECLDPPTTCISGN	2487

QY 485 -VDGDSGSGEPSPPTPGSTLTTPAVGLY 513
| : | : : | | | | |
Db 2488 CVNTPGSTYCDCCPPD---FELNPTRVGCV 2513

Search completed: May 8, 2002, 12:44:48
Job time: 349 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:39:24 ; Search time 17.04 Seconds

(without alignments)
1110.273 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916
Sequence: 1 MGVLYGALALAGLGPAP.....PSPFGSTLTPPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2912	99.9	575	1 TRBM_HUMAN	P07204 homo sapien
2	1899	65.1	577	1 TRBM_MOUSE	P15306 mus musculu
3	1092.5	37.5	356	1 TRBM_BOVIN	P06155 bos taurus
4	373	12.8	2907	1 FBN2_MOUSE	O61555 mus musculu
5	367.5	12.6	1184	1 FBN2_HUMAN	P98095 homo sapien
6	367	12.6	2871	1 FBN1_BOVIN	P98133 bos taurus
7	366	12.6	2871	1 FBN1_PIG	O91336 sus scrofa
8	365	12.5	2911	1 FBN2_HUMAN	P35556 homo sapien
9	363	12.4	2871	1 FBN1_HUMAN	P35555 homo sapien
10	361	12.4	2871	1 FBN1_MOUSE	O61554 mus musculu
11	358	12.3	956	1 MTN2_HUMAN	O00339 homo sapien
12	357	12.2	1221	1 FBN2_MOUSE	P37889 mus musculu
13	329.5	11.3	1964	1 MTN4_MOUSE	P31695 mus musculu
14	328.5	11.3	956	1 MTN2_MOUSE	O08746 mus musculu
15	327.5	11.2	1712	1 TGN2_MOUSE	O00918 ratu
16	324.5	11.1	712	1 FBN1_CAEBL	O77469 caenorhabd
17	320	11.0	1394	1 TGN2_HUMAN	P22064 homo sapien
18	315.5	10.8	703	1 FBN4_HUMAN	O95667 homo sapien
19	313	10.7	443	1 FBN4_HUMAN	P23142 homo sapien
20	308.5	10.6	443	1 FBN4_CRIGR	O55058 cricetulus
21	308	10.6	448	1 FBN5_HUMAN	O90495 homo sapien
22	308	10.6	705	1 FBN1_MOUSE	O08879 mus musculu
23	307	10.5	684	1 FBN1_CHICK	O73775 gallus gall
24	306.5	10.5	443	1 FBN4_MOUSE	O94779 mus musculu
25	302	10.4	448	1 FBN5_MOUSE	O94779 mus musculu
26	296	10.2	448	1 FBN5_MOUSE	O94779 mus musculu
27	296	10.2	2531	1 NTG1_MOUSE	O01705 mus musculu
28	288.5	9.9	493	1 FBN3_HUMAN	O12805 brachydanio
29	283.5	9.7	2437	1 NTG1_MOUSE	O07082 drosophila
30	282.5	9.7	2703	1 NTG1_MOUSE	O07082 mus musculu
31	280	9.6	2318	1 NTG1_MOUSE	O07082 mus musculu
32	277	9.5	2531	1 NTG1_MOUSE	O07082 mus musculu
33	276.5	9.5	493	1 FBN3_MOUSE	O35568 ratu

34	271	9.3	2444	1 NTG1_HUMAN	P46531 homo sapien
35	269.5	9.2	1217	1 EGF_MOUSE	P01132 mus musculu
36	263.5	9.0	619	1 MTN4_HUMAN	O95460 homo sapien
37	257.5	8.8	2524	1 NTG1_XENLA	P21783 xenopus lae
38	252.5	8.7	816	1 NPL2_RAT	O62918 ratu
39	252.5	8.7	835	1 CD97_HUMAN	P48960 homo sapien
40	251.5	8.6	816	1 NPL2_HUMAN	O99435 homo sapien
41	249.5	8.6	1133	1 EGF_RAT	P07522 ratu
42	248	8.5	624	1 MTN4_MOUSE	O89029 mus musculu
43	247	8.5	816	1 NPL2_MOUSE	O61220 mus musculu
44	247	8.5	816	1 NPL2_MOUSE	O90827 gallu
45	246	8.4	3051	1 NPL3_CAEBL	P34576 caenorhabd

ALIGNMENTS

RESULT	ID	TRBM_HUMAN	STANDARD:	PRT:	575 AA.
AC	P07204				
DT	01-APR-1988	(Rel. 07, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	THROMBOMODULIN PRECURSOR (FETOMODULIN) (TM) (CD141 ANTIGEN).				
GN	THBD OR TRBM.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88004395; PubMed=2820710;				
RA	Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I.,				
RA	Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.,				
RT	"Structure and expression of human thrombomodulin, a thrombin				
RT	receptor on endothelium acting as a cofactor for protein C				
RT	activation.";				
RL	EMBO J. 6:1891-1897(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88024950; PubMed=2822087;				
RA	Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.,				
RT	"Human thrombomodulin: complete cDNA sequence and chromosome				
RT	localization of the gene.";				
RL	Biochemistry 26:4350-4357(1987).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87317655; PubMed=2819876;				
RA	Jackson R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.,				
RT	"Human thrombomodulin gene is intron depleted: nucleic acid sequences				
RT	of the cDNA and gene predict protein structure and suggest sites of				
RT	regulatory control.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88227901; PubMed=2836377;				
RA	Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,				
RT	"Gene structure of human thrombomodulin, a cofactor for thrombin-				
RT	catalyzed activation of protein C.";				
RL	J. Biochem. 103:281-285(1988).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	Mathews L.,				
RL	Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.				
RX	MEDLINE=94029900; PubMed=8216207;				
RA	Gerleifer B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,				
RT	"Identification of the predominant glycosaminoglycan attachment site				
RT	in soluble recombinant human thrombomodulin: potential regulation of				

RT functional by glycosyltransferase competition for serine474.";
 RL Biochem. J. 295:131-140(1993).
 RP [7]
 RP STRUCTURE BY NMR OF 389-407.
 RX MEDLINE=96007474; PubMed=7559494;
 RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;
 RT "The structure of a 19-residue fragment from the C-loop of the fourth
 RT epidermal growth factor-like domain of thrombomodulin.";
 RL J. Biol. Chem. 270:23366-23372(1995).
 RN [8]
 RN STRUCTURE BY NMR OF 364-407.
 RX MEDLINE=96100636; PubMed=8528067;
 RA Weininger D.P., Hunter M.J., Komives E.A.;
 RT "Synthesis, activity, and preliminary structure of the fourth
 RT EGF-like domain of thrombomodulin.";
 RL Protein Sci. 4:1683-1695(1995).
 RN [9]
 RN STRUCTURE BY NMR OF 427-444.
 RX MEDLINE=95034791; PubMed=7947766;
 RA Srinivasan J., Hu S., Hirabai R., Zhu Y., Komives E.A., Ni F.;
 RT "Thrombin-bound structure of an EGF subdomain from human
 RT thrombomodulin determined by transferred nuclear Overhauser
 RT effects.";
 RL Biochemistry 33:13553-13560(1994).
 RN [10]
 RN STRUCTURE BY NMR OF 427-444.
 RX MEDLINE=96276211; PubMed=8745396;
 RA Hirabai R., Komives E.A., Ni F.;
 RT "Structural resiliency of an EGF-like subdomain bound to its target
 RT protein, thrombin.";
 RL Protein Sci. 5:195-203(1996).
 RN [11]
 RN STRUCTURE BY NMR OF 405-444.
 RX MEDLINE=98035729; PubMed=9367781;
 RA Sampol Benitez B.A., Hunter M.J., Weininger D.P., Komives E.A.;
 RT "Structure of the fifth EGF-like domain of thrombomodulin: an
 RT EGF-like domain with a novel disulfide-bonding pattern.";
 RL J. Mol. Biol. 273:913-926(1997).
 RN [12]
 RN VARIANT TED TYR-486.
 RX PubMed=7811989;
 RA Oehlin A.-K., Marlar R.A.;
 RT "The first mutation identified in the thrombomodulin gene in a
 RT 45-year-old man presenting with thromboembolic disease.";
 RL Blood 85:330-336(1995).
 RN [13]
 RN VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND I-501.
 RX PubMed=9198186;
 RA Oehlin A.-K., Norlund L., Marlar R.A.;
 RT "Thrombomodulin gene variations and thromboembolic disease.";
 RL Thromb. Haemost. 78:396-400(1997).
 RN [14]
 RN VARIANT VAL-473.
 RX PubMed=9157575;
 RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;
 RT "A common thrombomodulin amino acid dimorphism is associated with
 RT myocardial infarction.";
 RL Thromb. Haemost. 77:248-251(1997).
 RN [15]
 RN VARIANT THR-43.
 RX PubMed=9843165;
 RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,
 RT Stibbs P.J., Manger Cats V., Ireland H.;
 RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,
 RT and the risk of myocardial infarction in men.";
 RL Thromb. Haemost. 80:743-748(1998).
 RN [16]
 RN VARIANT VAL-473.
 RX PubMed=11245641;
 RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,
 RT Juneja H.;
 RT "Thrombomodulin Ala45Val polymorphism and risk of coronary heart
 RT disease.";

RL Circulation 103:1386-1389(2001).
 CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHelial CELL RECEPTOR
 CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
 CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
 CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
 CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
 CC FACTOR VA AND FACTOR VIII, AND THEREBY REDUCES THE AMOUNT OF
 CC THROMBIN GENERATED.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ENDOTHelial CELLS ARE UNIQUE IN SYNTHESIZING
 CC THROMBOMODULIN.
 CC -1- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED
 CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).
 CC -1- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,
 CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE
 CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE
 CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE
 CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.
 CC -1- SIMILARITY: CONTRAINS 6 EGF-LIKE DOMAINS.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD141 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X05495; CAA29045.1; -
 DR EMBL: M16552; AAB59508.1; -
 DR EMBL: J02973; AAB61175.1; -
 DR EMBL: D00210; BAA00149.1; -
 DR EMBL: AL049651; CAB51954.1; -
 DR PIR: A27073; A27073.
 DR PIR: A28307; A28307.
 DR PIR: A28680; A29680.
 DR PDB: 1EGT; 15-NOV-95.
 DR PDB: 1EGD; 20-JUN-96.
 DR PDB: 1EGE; 20-JUN-96.
 DR PDB: 1TMR; 08-JUN-95.
 DR PDB: 1ZAO; 29-JAN-96.
 DR PDB: 1ADX; 24-DEC-97.
 DR PDB: 2ADX; 24-DEC-97.
 DR GlycoSuiteDB: P07204; -
 DR MIM: 188040; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001491; Thmbomoduln.
 DR InterPro: IPR001304; Lectin_c.
 DR Pfam: PF00008; EGF; 5.
 DR Pfam: PF00059; lectin_c; 1.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_Like; 5.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR PROSITE: PS50041; C_Type_Lectin_2; 1.
 KW Endothelial cell; Receptor; blood coagulation; Repeat; Transmembrane;
 KW Glycoprotein; Signal; EGF-like domain; Disease mutation; Polymorphism;
 KW 3D-structure.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 575 THROMBOMODULIN.
 FT DOMAIN 22 515 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 516 539 POTENTIAL.
 FT DOMAIN 540 575 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 241 281 EGF-Like 1.
 FT DOMAIN 284 324 EGF-Like 2.

Query Match	65.1%;	Score 1839;	DB 1;	Length 577;
Best Local Similarity	65.5%;	Pred. No. 1.1e-117;		
Matches 341;	Conservative 46;	Mismatches 126;	Indels 8;	Gaps 4

RESULT	3
TRBM_BOVIN	
ID	TRBM_BOVIN
AC	P06579;
DT	01-JAN-1988 (rel. 06, Created)
DT	01-JAN-1988 (rel. 06, Last sequence update)
DT	20-JAN-2001 (rel. 40, Last annotation update)
DE	THROMBOMODULIN (FETOMODULIN) (TM) (FRAGMENT).
GN	THBD.
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=87067408; PubMed=3024152;
RA	Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;
RT	"characterization of a thrombomodulin cDNA reveals structural
RT	similarity to the low density lipoprotein receptor.";
RL	Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).
CC	- FUNCTION: THROMBOMULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC	THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC	COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC	ACTIVATED PROTEIN C (PROTEIN CA). ONCE ENOLAYD, PROTEIN CA
CC	SECTIONS THE ACTIVATED COPAFORS OF THE COAGULATION MECHANISM,
CC	FACTOR VA AND FACTOR VIII, AND THEREBY REDUCES THE AMOUNT OF

CC	-1-	THROMBIN GENERATED.
CC	-1-	SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC	-1-	TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING THROMBOMODULIN.
CC	-1-	SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC		-----
DR	PMB;	M14657; AAA30785.1; -.
DR	PIR;	A25918; A25918.
DR	HSSP;	P07204; 1EGF.
DR	InterPro;	IPRO00152; Asx_hydroxyl.
DR	InterPro;	IPIRO00561; EGF-like.
DR	InterPro;	IPIRO01861; EGF_Ca.
DR	Pfam;	PF00008; EGF_5.
DR	SMART;	SM00179; EGF_CA; 1.
DR	SMART;	SM00001; EGF_Like; 3.
DR	PROSITE;	PS00010; ASX_HYDROXYL; 2.
DR	PROSITE;	PS00022; EGF_1; FALSE_NNG.
DR	PROSITE;	PS01186; EGF_2; 3.
DR	PROSITE;	PS01187; EGF_CA; 2.
KW	Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane; Glycoprotein; EGF-like domain.	
FT	NON_TER	1 1
FT	DOMAIN	1 296 EXTRACELLULAR (POTENTIAL) .
FT	TRANSEM	297 320 POTENTIAL.
FT	DOMAIN	321 356 CYTOPLASMIC (POTENTIAL) .
FT	DOMAIN	17 57 EGF-LIKE 1.
FT	DOMAIN	60 98 EGF-LIKE 2.
FT	DOMAIN	99 137 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL) .
FT	DOMAIN	139 179 EGF-LIKE 4.
FT	DOMAIN	178 214 EGF-LIKE 5.
FT	DOMAIN	215 254 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL) .
FT	DISULEID	21 32 BY SIMILARITY.
FT	DISULEID	28 41 BY SIMILARITY.
FT	DISULEID	43 56 BY SIMILARITY.
FT	DISULEID	64 72 BY SIMILARITY.
FT	DISULEID	68 82 BY SIMILARITY.
FT	DISULEID	84 97 BY SIMILARITY.
FT	DISULEID	103 114 BY SIMILARITY.
FT	DISULEID	110 123 BY SIMILARITY.
FT	DISULEID	125 136 BY SIMILARITY.
FT	DISULEID	143 152 BY SIMILARITY.
FT	DISULEID	148 162 BY SIMILARITY.
FT	DISULEID	164 178 BY SIMILARITY.
FT	DISULEID	182 191 BY SIMILARITY.
FT	DISULEID	187 199 BY SIMILARITY.
FT	DISULEID	201 213 BY SIMILARITY.
FT	DISULEID	219 228 BY SIMILARITY.
FT	DISULEID	224 237 BY SIMILARITY.
FT	DISULEID	239 253 BY SIMILARITY.
FT	CARBONYD	271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY) .
QO	SEQUENCE	356 AA; 37795 MW; 2394A1E097ABE1093 CRC64; 29B44F097ABE1093 CRC64;

[illegible]

FT	DISULFID	329	343	BY SIMILARITY.
FT	DISULFID	345	358	BY SIMILARITY.
FT	DISULFID	491	503	BY SIMILARITY.
FT	DISULFID	498	512	BY SIMILARITY.
FT	DISULFID	514	526	BY SIMILARITY.
FT	DISULFID	532	542	BY SIMILARITY.
FT	DISULFID	537	551	BY SIMILARITY.
FT	DISULFID	553	566	BY SIMILARITY.
FT	DISULFID	572	584	BY SIMILARITY.
FT	DISULFID	579	593	BY SIMILARITY.
FT	DISULFID	595	608	BY SIMILARITY.
FT	DISULFID	614	625	BY SIMILARITY.
FT	DISULFID	620	634	BY SIMILARITY.
FT	DISULFID	636	649	BY SIMILARITY.
FT	DISULFID	655	666	BY SIMILARITY.
FT	DISULFID	661	675	BY SIMILARITY.
FT	DISULFID	677	690	BY SIMILARITY.
FT	DISULFID	765	777	BY SIMILARITY.
FT	DISULFID	772	786	BY SIMILARITY.
FT	DISULFID	788	801	BY SIMILARITY.
FT	DISULFID	807	819	BY SIMILARITY.
FT	DISULFID	814	828	BY SIMILARITY.
FT	DISULFID	830	843	BY SIMILARITY.
FT	DISULFID	849	859	BY SIMILARITY.
FT	DISULFID	854	868	BY SIMILARITY.
FT	DISULFID	870	883	BY SIMILARITY.
FT	DISULFID	952	964	BY SIMILARITY.
FT	DISULFID	959	973	BY SIMILARITY.
FT	DISULFID	975	988	BY SIMILARITY.
FT	DISULFID	1070	1082	BY SIMILARITY.
FT	DISULFID	1077	1091	BY SIMILARITY.
FT	DISULFID	1093	1106	BY SIMILARITY.
FT	DISULFID	1112	1124	BY SIMILARITY.
FT	DISULFID	1119	1133	BY SIMILARITY.
FT	DISULFID	1135	1149	BY SIMILARITY.
FT	DISULFID	1155	1167	BY SIMILARITY.
FT	DISULFID	1162	1176	BY SIMILARITY.
FT	DISULFID	1178	1191	BY SIMILARITY.
FT	DISULFID	1197	1209	BY SIMILARITY.
FT	DISULFID	1204	1218	BY SIMILARITY.
FT	DISULFID	1220	1233	BY SIMILARITY.
FT	DISULFID	1239	1250	BY SIMILARITY.
FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DISULFID	1261	1274	BY SIMILARITY.
FT	DISULFID	1280	1292	BY SIMILARITY.
FT	DISULFID	1287	1301	BY SIMILARITY.
FT	DISULFID	1303	1316	BY SIMILARITY.
FT	DISULFID	1322	1334	BY SIMILARITY.
FT	DISULFID	1329	1343	BY SIMILARITY.
FT	DISULFID	1345	1358	BY SIMILARITY.
FT	DISULFID	1364	1377	BY SIMILARITY.
FT	DISULFID	1371	1386	BY SIMILARITY.
FT	DISULFID	1386	1399	BY SIMILARITY.
FT	DISULFID	1405	1418	BY SIMILARITY.
FT	DISULFID	1412	1427	BY SIMILARITY.
FT	DISULFID	1429	1440	BY SIMILARITY.
FT	DISULFID	1446	1458	BY SIMILARITY.
FT	DISULFID	1453	1467	BY SIMILARITY.
FT	DISULFID	1469	1482	BY SIMILARITY.
FT	DISULFID	1488	1499	BY SIMILARITY.
FT	DISULFID	1494	1508	BY SIMILARITY.
FT	DISULFID	1510	1523	BY SIMILARITY.
FT	DISULFID	1529	1540	BY SIMILARITY.
FT	DISULFID	1535	1549	BY SIMILARITY.
FT	DISULFID	1551	1564	BY SIMILARITY.
FT	DISULFID	1647	1659	BY SIMILARITY.
FT	DISULFID	1654	1668	BY SIMILARITY.
FT	DISULFID	1670	1683	BY SIMILARITY.
FT	DISULFID	1689	1701	BY SIMILARITY.
FT	DISULFID	1696	1710	BY SIMILARITY.

Query Match 12.88; Score 373; DB 1; Length 2907;
Best Local Similarity 32.98; Pred. No. 4, 5e-17;

Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;			
QY	244	DCSEVNGCGHAACNAPGAPROCCPAGALQADGSCGASATQSCNDCEHPCVNPDP--	301
Db	1238	ECMIMGCCDTCTGNSGSEGCSEGVAMPDGSCTA-----DIDE--CENNPIC	1287
QY	302	-----QPSYSQMCETGYRLADQRCEDVDCEILEPSPCP-ORCVNTOGGECHCYP	353
Db	1288	DGGCCNIPPEYRCLCYDGFMAKMKICIDVNEEDLNPNICMFECENTGSGFICHOQL	1347
QY	354	NYDLVGE--CVPPVPC--FRANCYQCQPLN-QTSYLCVCAEFA-----PIP	398
Db	1348	GYSVKKGTTCCTD-VDECEIGAHNCMDHASCINVPSEFCSCREGWNGIKICIDLECA	1406
QY	399	HEPRCOMFCNACPADCDPTOAS--CECEGYILDDGFICTDIDE-----CENG-	448
Db	1407	NGTHQCSI-----NAQC-VNTPGSGYRCACSEGF-TGGGFCSDYDCAENTNCENGQ	1457
QY	449	-----GF-----GS-----GCNHLPGTFECTICGDSALV	473
Db	1458	CLNYPGAYRCECEMGFTPASDRSCODIDECGSONICVGTGCTNMDPGFHICIDGVELD	1517
QY	474	RHIG--TDCD	481
Db	1518	RTGSGNCTDID	1527
RESULT 5			
FBUL2_HUMAN	STANDARD;	PRT:	1184 AA.
AC	P80055;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	FBULIN-2 PRECURSOR.		
GN	FBUL2.		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain; Pubmed=7806230;		
RX	MEDLINE=95104855; Pubmed=7806230;		
RA	Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R.,		
RA	Chu M.-L.,		
RT	"Fibulin-2 (FBUL2): human cDNA sequence, mRNA expression, and mapping		
RT	of the gene on human and mouse chromosomes";		
RL	Genomics 22:425-430(1994).		
CC	-1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS		
CC	CALCIUM DEPENDENT.		
CC	-1- SUBUNIT: HOMODIMER: DISULFIDE-LINKED.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH		
CC	BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.		
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.		
CC	-1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; X82494; CA57876.1; -		
DR	HSP; P07204; IFGD.		
DR	MM; 135821; -		
DR	InterPro; IPR000020; Anaphylatoxin.		
DR	InterPro; IPR000152; Asx hydroxyl.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR001881; EGF_Ca.		


```

CC      EGF-LIKE DOMAINS.
CC      -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L28748; AAT4122.1; -.
DR      HSSP: P35555; IAPJ.
DR      InterPro: IPR000152; Asx_hydroxyl.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR001438; EGF-II.
DR      InterPro: IPR002212; TB.
DR      Pfam: PF00008; EGF; 46.
DR      Pfam: PF00683; TB; 9.
DR      PRINTS: PR00010; EGFBL00D.
DR      SMART: SM00179; EGF_CA_42.
DR      SMART: SM00001; EGF_Like; 4.
DR      PROSITE: PS00010; ASX_HYDROXYL; 43.
DR      PROSITE: PS00022; EGF_1; 2.
DR      PROSITE: PS01186; EGF_2; 38.
DR      PROSITE: PS01187; EGF_CA; 45.
KW      Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW      Repeat; Signal; Multigene family.
FT      SIGNAL 1 27
FT      CHAIN 28 2871
FT      DOMAIN 81 112 FIBRILLIN 1.
FT      DOMAIN 115 146 EGF-LIKE 1, NON-CALCIUM BINDING.
FT      DOMAIN 147 178 EGF-LIKE 2, NON-CALCIUM BINDING.
FT      DOMAIN 246 287 EGF-LIKE 3, NON-CALCIUM BINDING.
FT      DOMAIN 288 329 EGF-LIKE 4, CALCIUM-BINDING.
FT      REPEAT 330 390 EGF-LIKE 5, CALCIUM-BINDING.
FT      DOMAIN 332 390 TGFBRP 1.
FT      DOMAIN 392 446 PRO-RICH.
FT      DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT      DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT      DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT      DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT      DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT      REPEAT 654 722 TGFBRP 2.
FT      DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT      DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT      DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT      DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT      REPEAT 952 1027 TGFBRP 3.
FT      DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT      DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT      DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT      DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT      DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT      DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT      DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT      DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT      DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT      DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT      DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT      DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT      REPEAT 1528 1605 TGFBRP 4.
FT      DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT      DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT      REPEAT 1689 1765 TGFBRP 5.
FT      DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT      DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT      DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT      DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT      DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT      DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT      DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT      REPEAT 2055 2126 TGFBRP 6.

```

```

FT      DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT      DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT      DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT      DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT      REPEAT 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT      DOMAIN 2333 2402 TGFBRP 7.
FT      DOMAIN 2403 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT      DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT      DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT      DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT      DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT      DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT      DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT      DISULFID 85 94 BY SIMILARITY.
FT      DISULFID 89 100 BY SIMILARITY.
FT      DISULFID 102 111 BY SIMILARITY.
FT      DISULFID 113 129 BY SIMILARITY.
FT      DISULFID 123 134 BY SIMILARITY.
FT      DISULFID 136 145 BY SIMILARITY.
FT      DISULFID 150 160 BY SIMILARITY.
FT      DISULFID 154 166 BY SIMILARITY.
FT      DISULFID 168 177 BY SIMILARITY.
FT      DISULFID 180 262 BY SIMILARITY.
FT      DISULFID 250 271 BY SIMILARITY.
FT      DISULFID 273 286 BY SIMILARITY.
FT      DISULFID 292 304 BY SIMILARITY.
FT      DISULFID 299 313 BY SIMILARITY.
FT      DISULFID 315 328 BY SIMILARITY.
FT      DISULFID 453 465 BY SIMILARITY.
FT      DISULFID 460 474 BY SIMILARITY.
FT      DISULFID 476 488 BY SIMILARITY.
FT      DISULFID 494 504 BY SIMILARITY.
FT      DISULFID 499 513 BY SIMILARITY.
FT      DISULFID 515 528 BY SIMILARITY.
FT      DISULFID 534 546 BY SIMILARITY.
FT      DISULFID 541 555 BY SIMILARITY.
FT      DISULFID 557 570 BY SIMILARITY.
FT      DISULFID 576 587 BY SIMILARITY.
FT      DISULFID 582 596 BY SIMILARITY.
FT      DISULFID 598 611 BY SIMILARITY.
FT      DISULFID 617 628 BY SIMILARITY.
FT      DISULFID 623 637 BY SIMILARITY.
FT      DISULFID 639 652 BY SIMILARITY.
FT      DISULFID 727 735 BY SIMILARITY.
FT      DISULFID 734 748 BY SIMILARITY.
FT      DISULFID 750 763 BY SIMILARITY.
FT      DISULFID 769 781 BY SIMILARITY.
FT      DISULFID 776 790 BY SIMILARITY.
FT      DISULFID 792 805 BY SIMILARITY.
FT      DISULFID 811 821 BY SIMILARITY.
FT      DISULFID 816 830 BY SIMILARITY.
FT      DISULFID 832 845 BY SIMILARITY.
FT      DISULFID 914 926 BY SIMILARITY.
FT      DISULFID 921 935 BY SIMILARITY.
FT      DISULFID 937 950 BY SIMILARITY.
FT      DISULFID 1032 1044 BY SIMILARITY.
FT      DISULFID 1039 1053 BY SIMILARITY.
FT      DISULFID 1055 1068 BY SIMILARITY.
FT      DISULFID 1074 1086 BY SIMILARITY.
FT      DISULFID 1081 1095 BY SIMILARITY.
FT      DISULFID 1097 1111 BY SIMILARITY.
FT      DISULFID 1117 1129 BY SIMILARITY.
FT      DISULFID 1124 1138 BY SIMILARITY.
FT      DISULFID 1140 1153 BY SIMILARITY.
FT      DISULFID 1159 1171 BY SIMILARITY.
FT      DISULFID 1166 1180 BY SIMILARITY.
FT      DISULFID 1182 1195 BY SIMILARITY.
FT      DISULFID 1201 1212 BY SIMILARITY.
FT      DISULFID 1208 1221 BY SIMILARITY.
FT      DISULFID 1223 1236 BY SIMILARITY.
FT      DISULFID 1242 1254 BY SIMILARITY.
FT      DISULFID 1249 1263 BY SIMILARITY.
FT      DISULFID 1265 1278 BY SIMILARITY.

```

Query Match	Best Local Similarity	12.6%	Score 367;	DB 1;	Length 2871;
Matches 111;	Conservative	26;	Mismatches 109;	Indels 106;	Gaps 20;
FT DISULFID 1284	BY SIMILARITY.				
FT DISULFID 1291	BY SIMILARITY.				
FT DISULFID 1307	BY SIMILARITY.				
FT DISULFID 1326	BY SIMILARITY.				
FT DISULFID 1333	BY SIMILARITY.				
FT DISULFID 1350	BY SIMILARITY.				
FT DISULFID 1367	BY SIMILARITY.				
FT DISULFID 1374	BY SIMILARITY.				
FT DISULFID 1391	BY SIMILARITY.				
FT DISULFID 1408	BY SIMILARITY.				
FT DISULFID 1415	BY SIMILARITY.				
FT DISULFID 1431	BY SIMILARITY.				
FT DISULFID 1450	BY SIMILARITY.				
FT DISULFID 1456	BY SIMILARITY.				
FT DISULFID 1472	BY SIMILARITY.				
FT DISULFID 1491	BY SIMILARITY.				
FT DISULFID 1497	BY SIMILARITY.				
FT DISULFID 1513	BY SIMILARITY.				
FT DISULFID 1610	BY SIMILARITY.				
Query Match	Best Local Similarity	12.6%;	Score 367;	DB 1;	Length 2871;
Matches 111;	Conservative	26;	Mismatches 109;	Indels 106;	Gaps 20;
OY 244 DCSVNGGCEHACNAIPGAPRCQCPAGALQADGNSCTASATQSCNDLCEHFCVPPNP---	300				
Db 1200 ECSINMGCEFFCTSTSEGSSEGCSCPGFALMPDRSCT-----DIDE--CEDNPNIC	1249				
OY 301 -----DQGSASCCEGUYRLAQAQHNREVDQDCLLEPSPC--PQRCVTOGGGFECHCP	353				
Db 1250 DGGCCTNIPGEFRCLCYDGFMASEDMKTCVYVNECDLNPITLSTGICENTKSFICHCDM	1309				
OY 354 NYDLVGE--CYEPVDP--FRANCEYOCQPLNQ--SYLCAEAGFA-----PIP	398				
Db 1310 GYSGRKGTGCTD--INECIGAHNCDRHAVCTNTAGSFRCSCSPGWIGDIGIKCTDLDECS	1368				
OY 399 HEPHRCQMFNCTACRACDDPTQMS--CECEGYTLDDGFICTDIDEC--EN-----	447				
Db 1369 NQTHMCSQH-----ADC-KMTMGSYRCLCKEY--TGDFCTCTDLDECSNMLNCGNQ	1419				
OY 448 -----GGF-----CS-----GYCHNLPGTFECLCGPSDALY	473				
Db 1420 CLNARGYRCECDMGFVPSADGKACEDIDECSLPNICVGTGHNLPLGLRCECEIYEILD	1479				
OY 474 RHIG-----TDCDSGK--VDGGSSECEPPSPPTGSLTPPAPVGLY	513				
Db 1480 RSGGCTDVNECLDPTTCISGNCVMTPGSYTCDCPDP---PELNPTRVGCV	1527				
RESULT 7					
FBNI_PIG	STANDARD:	PRT;	2871	AA.	
AC 09TV36;					
DR 20-AUG-2001 (Rel. 40, Created)					
DT 20-AUG-2001 (Rel. 40, Last sequence update)					
DT 20-AUG-2001 (Rel. 40, Last annotation update)					
DE FIBRILLIN 1 PRECURSOR.					
GN FBNI.					
OS Sus scrofa (Pig).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.					
OX NCBI_TaxID=9823;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=Lung;					
RX MEDLINE=99156858; PubMed=10036187;					
RA Blevy N.J., Eldadah Z.A., Moore C.S., Stetten G., Spencer F.,					
RA Dietz H.C.;					
RT "Revised genomic organization of FBNI and significance for regulated					
RT gene expression";					
RT Genomics 56:70-77(1999).					
CC -! FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS					
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE					

CC	LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC	-1- PMM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC	FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC	MICROFIBRILS.
CC	-1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC	EGF-LIKE DOMAINS.
CC	-1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AF073800; ADB50328.1; -.
DR	HSP; P35555; IAPJ.
DR	InterPro: IPRO00152; Asx_hydroxyl.
DR	InterPro: IPRO00561; EGF-like.
DR	InterPro: IPRO01881; EGF Ca.
DR	InterPro: IPRO01438; EGF-II.
DR	InterPro: IPRO02212; Znf-C2H2.
DR	InterPro: IPRO00822; Znf-C2H2.
DR	Pfam; PF00683; TB; 9.
DR	Pfam; PF00683; TB; 9.
DR	PRINTS; PR0010; EGFBL0D.
DR	SMART; SMO0179; EGF_CA; 40.
DR	SMART; SMO0001; EGF_Like; 6.
DR	PROSITE; PS00010; ASX_HYDROXYL_41.
DR	PROSITE; PS00022; EGF_L1; UNKNOWN_2.
DR	PROSITE; PS01186; EGF_2; 36.
KW	Extracellular matrix; Calcium-binding; glycoprotein; EGF-like domain;
KM	Repeat; signal; Multigene family.
FT	SIGNAL 1 27 POTENTIAL.
FT	CHAIN 28 2871 FIBRILLIN 1.
FT	DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT	DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT	DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT	DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT	DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT	REPEAT 330 390 TGFBP 1.
FT	DOMAIN 392 446 PRO-RICH.
FT	DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT	DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT	DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT	DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT	DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT	REPEAT 654 722 TGFBP 2.
FT	DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT	DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT	DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT	DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT	REPEAT 952 1027 TGFBP 3.
FT	DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT	DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT	DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT	DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT	DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT	DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT	DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT	DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT	DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT	DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT	DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT	DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT	REPEAT 1528 1605 TGFBP 4.
FT	DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT	DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT	REPEAT 1689 1765 TGFBP 5.
FT	DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT	DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.

```
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2055 2126 TGFBP 6.
FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2333 2401 TGFBP 7.
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2524 2563 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 177 262 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 513 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 633 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 821 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 845 855 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
```

```
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
FT DISULFID 1770 1782 BY SIMILARITY.
FT DISULFID 1777 1791 BY SIMILARITY.
FT DISULFID 1793 1806 BY SIMILARITY.
FT DISULFID 1812 1824 BY SIMILARITY.
FT DISULFID 1818 1833 BY SIMILARITY.
```

Query Match 12.6%: Score 366; DB 1; Length 2871;

Best Local Similarity 30.6%: Pred No. 1,3e-16; Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

```
QY 244 DCSVENGCEHACNAPALGAPRCQCPAGALQADGRSTASATQSCNDLCEHFCVBNP--- 300
      :||: ||||| :| :| ||| ||||| :| :| |||
Db 1200 ECSIMNGCCFTCTNBSGSEYSCQPGFALMPQRCST-----DIDE--CEDNPNIC 1249

QY 301 -----DQPSYSCMCEYGRILADQHRCEVDVDCILFESPC--PQRCVNTQSGFECGYR 353
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1250 DGGCTNIPPEYKCLCYDGMASEDMKTCVDNECDLNNICISGTCENTKSGFYCHCDM 1309

QY 354 NYDLVDE--CVRPVDP--FRANCEYOCQPLNOT--SYLCAEGRA-----PIP 398
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1310 GYSKKKKTGCTD- INCEIGAHNCDRHAVCTYATGAFSCSPGWIQGIKCTDLDEC 1368

QY 399 HEPRRCMPFCNOTACPADCDPNTQAS--CECPGGYILDDGFTCTDIDEC--EN----- 447
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1369 NGTHMCSQH-----ADC-KNTMGSYRCICKEGY--TGDGFTCADLDECSENVKLCGNVQ 1419

QY 448 -----GGF-----CS-----GVCHNLPGRFECICGPDALV 473
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1420 CLVAPGSYHCEYDMGVPFSPADRKSCVDSDECSLPINLCVGTCHNLPGLFRCC----- 1472

QY 474 RHIGTDCDSKVDGSDSGSGBPSPS-----PTPGS-----TLTPPAVGV 513
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1473 -ELIGYELDRSGNCTDVNECTLEPPTCISGNCVYTPGSYCVCPPELMPTRVGCY 1527

RESULT 8
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBRILLIN 2 PRECURSOR.
GN FBN2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


DR	PRINTS: PR00010; EGF_CA: 43.
DR	SMART: SM000179; EGF_LIKE: 43.
DR	SMART: SM00001; EGF_LIKE: 3.
DR	PROSITE: PS00010; ASX_HYDROXYL: 43.
DR	PROSITE: PS00022; EGF_1: 2.
DR	PROSITE: PS01186; EGF_2; 37.
DR	PROSITE: PS01187; EGF_CA; 43.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Signal; Multigene family; Disease mutation; Polymorphism.
FT	SIGNAL
FT	1
FT	28
FT	POTENTIAL.
FT	1
FT	FIBRILLIN 2.
FT	EGF-LIKE 1, NON-CALCIUM BINDING.
FT	111 142
FT	EGF-LIKE 2, NON-CALCIUM BINDING.
FT	145 176
FT	EGF-LIKE 3, NON-CALCIUM BINDING.
FT	176 207
FT	EGF-LIKE 4, CALCIUM-BINDING.
FT	215 316
FT	EGF-LIKE 5, CALCIUM-BINDING.
FT	317 358
FT	REPEAT
FT	359 425
FT	EGF-LIKE 6, NON-CALCIUM BINDING.
FT	493 533
FT	EGF-LIKE 7, CALCIUM-BINDING.
FT	534 573
FT	EGF-LIKE 8, CALCIUM-BINDING.
FT	574 615
FT	EGF-LIKE 9, CALCIUM-BINDING.
FT	616 656
FT	EGF-LIKE 10, CALCIUM-BINDING.
FT	657 697
FT	REPEAT
FT	698 766
FT	EGF-LIKE 11, CALCIUM-BINDING.
FT	767 808
FT	EGF-LIKE 12, CALCIUM-BINDING.
FT	809 850
FT	EGF-LIKE 13, CALCIUM-BINDING.
FT	851 995
FT	EGF-LIKE 14, CALCIUM-BINDING.
FT	954 995
FT	REPEAT
FT	996 1071
FT	EGF-LIKE 15, CALCIUM-BINDING.
FT	1072 1113
FT	EGF-LIKE 16, CALCIUM-BINDING.
FT	1114 1156
FT	EGF-LIKE 17, CALCIUM-BINDING.
FT	1157 1198
FT	EGF-LIKE 18, CALCIUM-BINDING.
FT	1199 1240
FT	EGF-LIKE 19, CALCIUM-BINDING.
FT	1241 1281
FT	EGF-LIKE 20, CALCIUM-BINDING.
FT	1282 1323
FT	EGF-LIKE 21, CALCIUM-BINDING.
FT	1324 1365
FT	EGF-LIKE 22, CALCIUM-BINDING.
FT	1366 1406
FT	EGF-LIKE 23, CALCIUM-BINDING.
FT	1407 1447
FT	EGF-LIKE 24, CALCIUM-BINDING.
FT	1448 1489
FT	EGF-LIKE 25, CALCIUM-BINDING.
FT	1490 1530
FT	EGF-LIKE 26, CALCIUM-BINDING.
FT	1531 1571
FT	REPEAT
FT	1572 1648
FT	EGF-LIKE 27, CALCIUM-BINDING.
FT	1649 1690
FT	EGF-LIKE 28, CALCIUM-BINDING.
FT	1691 1732
FT	REPEAT
FT	1733 1806
FT	EGF-LIKE 29, CALCIUM-BINDING.
FT	1807 1848
FT	EGF-LIKE 30, CALCIUM-BINDING.
FT	1849 1890
FT	EGF-LIKE 31, CALCIUM-BINDING.
FT	1891 1932
FT	EGF-LIKE 32, CALCIUM-BINDING.
FT	1933 1971
FT	EGF-LIKE 33, CALCIUM-BINDING.
FT	1972 2014
FT	EGF-LIKE 34, CALCIUM-BINDING.
FT	2015 2054
FT	EGF-LIKE 35, CALCIUM-BINDING.
FT	2055 2096
FT	REPEAT
FT	2097 2169
FT	EGF-LIKE 36, CALCIUM-BINDING.
FT	2170 2211
FT	EGF-LIKE 37, CALCIUM-BINDING.
FT	2212 2251
FT	EGF-LIKE 38, CALCIUM-BINDING.
FT	2252 2292
FT	EGF-LIKE 39, CALCIUM-BINDING.
FT	2293 2336
FT	EGF-LIKE 40, CALCIUM-BINDING.
FT	2337 2378
FT	REPEAT
FT	2379 2447
FT	EGF-LIKE 41, CALCIUM-BINDING.
FT	2448 2489
FT	EGF-LIKE 42, CALCIUM-BINDING.
FT	2490 2530
FT	EGF-LIKE 43, CALCIUM-BINDING.
FT	2531 2569
FT	EGF-LIKE 44, CALCIUM-BINDING.
FT	2570 2612
FT	EGF-LIKE 45, CALCIUM-BINDING.
FT	2613 2652
FT	EGF-LIKE 46, CALCIUM-BINDING.
FT	2653 2693
FT	EGF-LIKE 47, CALCIUM-BINDING.
FT	2694 2733
FT	DOMAIN
FT	115 124
FT	DISULFID
FT	119 130
FT	DISULFID
FT	132 141
FT	DISULFID
FT	143 159
FT	DISULFID
FT	153 164
FT	DISULFID
FT	166 175
FT	DISULFID
FT	180 190
FT	DISULFID
FT	184 195

FT	DISULFID	197	206	BY SIMILARITY.
FT	DISULFID	279	281	BY SIMILARITY.
FT	DISULFID	286	300	BY SIMILARITY.
FT	DISULFID	302	315	BY SIMILARITY.
FT	DISULFID	321	333	BY SIMILARITY.
FT	DISULFID	328	342	BY SIMILARITY.
FT	DISULFID	344	357	BY SIMILARITY.
FT	DISULFID	437	509	BY SIMILARITY.
FT	DISULFID	504	518	BY SIMILARITY.
FT	DISULFID	520	532	BY SIMILARITY.
FT	DISULFID	538	548	BY SIMILARITY.
FT	DISULFID	543	557	BY SIMILARITY.
FT	DISULFID	559	572	BY SIMILARITY.
FT	DISULFID	578	590	BY SIMILARITY.
FT	DISULFID	585	599	BY SIMILARITY.
FT	DISULFID	601	614	BY SIMILARITY.
FT	DISULFID	620	631	BY SIMILARITY.
FT	DISULFID	626	640	BY SIMILARITY.
FT	DISULFID	642	655	BY SIMILARITY.
FT	DISULFID	661	672	BY SIMILARITY.
FT	DISULFID	667	681	BY SIMILARITY.
FT	DISULFID	683	696	BY SIMILARITY.
FT	DISULFID	771	783	BY SIMILARITY.
FT	DISULFID	778	792	BY SIMILARITY.
FT	DISULFID	794	807	BY SIMILARITY.
FT	DISULFID	813	825	BY SIMILARITY.
FT	DISULFID	820	834	BY SIMILARITY.
FT	DISULFID	836	849	BY SIMILARITY.
FT	DISULFID	855	865	BY SIMILARITY.
FT	DISULFID	860	874	BY SIMILARITY.
FT	DISULFID	876	889	BY SIMILARITY.
FT	DISULFID	958	970	BY SIMILARITY.
FT	DISULFID	965	979	BY SIMILARITY.
FT	DISULFID	981	994	BY SIMILARITY.
FT	DISULFID	1076	1088	BY SIMILARITY.
FT	DISULFID	1083	1097	BY SIMILARITY.
FT	DISULFID	1099	1112	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1139	BY SIMILARITY.
FT	DISULFID	1141	1155	BY SIMILARITY.
FT	DISULFID	1161	1173	BY SIMILARITY.
FT	DISULFID	1168	1182	BY SIMILARITY.
FT	DISULFID	1184	1197	BY SIMILARITY.
FT	DISULFID	1203	1215	BY SIMILARITY.
FT	DISULFID	1210	1224	BY SIMILARITY.

Query Match 12.5% Score 365; DB 1; Length 2911;
 Best Local Similarity 33.1%; Pred. No. 1.5e-16;
 Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

OY	244	DCSVENGCEHACNAIPGAPROCCPAGALQADGSRCTASATQSCNDLCHEFCVNPDP--	301
DB	1244	ECMINMGGDYQCTSESGYSCSGEVALMPDGRSCA-----DIDE--CENNPDI	1293
OY	302	-----OPGSYSCMETGYRLAADOHCEVDPCILERSPCP-QRCVNTGSGFCHCYP	353
DB	1294	DGGQCTNIGETRCICLCYGFAMSMOKTICIDNEDLNSNIMCEPCENTKGSFICHQOL	1353
OY	354	NDLVDFE--CVEPVDPC--FRANCEYOQOPLN-OTSLVCVACGAFAPR-----EPHR	403
DB	1354	GYSVKKGTGCTD-VDECEIGAHNCMDHASCINIGSFSCREGW--IGNIKICIDLE	1410
OY	404	COMFCQTRACPDADPNQAS--CECPBGYIILDGFICTIDDE-----CENG-----	448
DB	1411	CSNGTHQCSINAG--VNPGRYKACSEGF-TGDEFTSDVDECAENINLCEMGQCLNVP	1468
OY	449	-----GF-----CSGVCHLPTGTEFCICIGPPSALVRHIG-	477
DB	1469	GAIRCECEGFTPADSRSCQDIDECSPONICVSGTCNMLPMGFHCTDDGYELDRTSGN	1528
OY	478	-TDCD 481	
DB	1529	CMDID 1533	

RESULT	9
FBN1_HUMAN	STANDARD; PRT; 2871 AA.
ID	FBN1_HUMAN
AC	P35555; 01-JUN-1994 (Rel. 29, Created)
DT	01-JUN-1994 (Rel. 29, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	FIBRILLIN 1 PRECURSOR.
GN	FBN1 OR FBN.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX	NCBI_TaxID=9606;
ON	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Placenta;
RX	MEDLINE=93372860; PubMed=8364578;
RA	Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,
RT	Pangilinan T., Bonadio J.;
RT	"Genomic organization of the sequence coding for fibrillin, the
RL	defective gene product in Marfan syndrome.";
RN	Hum. Mol. Genet. 2:961-968(1993).
RP	SEQUENCE OF 1-932 FROM N.A.
RC	TISSUE=Placenta; and Fibroblast;
RX	MEDLINE=94010947; PubMed=7691719;
RA	Corson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;
RT	"Fibrillin binds calcium and is coded by cDNAs that reveal a
RT	multidomain structure and alternatively spliced exons at the 5'
RT	end.";
RL	Genomics 17:476-484(1993).
RP	SEQUENCE OF 899-2871 FROM N.A.
RX	MEDLINE=91304568; PubMed=1852207;
RA	Maslen C.L., Corson G.M., Maddox B.K., Gianvillie R.W., Sakai L.Y.;
RT	"Partial sequence of a candidate gene for the Marfan syndrome.";
RL	Nature 352:334-337(1991).
RP	SEQUENCE OF 813-1313 FROM N.A.
RX	MEDLINE=91304567; PubMed=1852206;
RA	Lee B., Godfrey M., Vitale E., Hort H., Mattei M.-G., Sartarazi M.,
RT	Tsipouras P., Ramirez F., Hollister D.W.;
RT	"Linkage of Marfan syndrome and a phenotypically related disorder to
RL	two different fibrillin genes.";
RL	Nature 352:330-334(1991).
RP	CHARACTERIZATION.
RX	MEDLINE=91317849; PubMed=1860873;
RA	Sakai L.Y., Keene D.R., Gianvillie R.W., Bachinger H.P.;
RT	"Purification and partial characterization of fibrillin, a cysteine-
RT	rich structural component of connective tissue microfibrils.";
RL	J. Biol. Chem. 266:14763-14770(1991).
RP	STRUCTURE BY NMR OF 2054-2125.
RX	MEDLINE=98031893; PubMed=9362480;
RA	Yuan X., Downing A.K., Knott V., Handford P.A.;
RT	"Solution structure of the transforming growth factor beta-binding
RT	protein-like module, a domain associated with matrix fibrils.";
RL	EMBO J. 16:6659-6666(1997).
RP	STRUCTURE BY NMR OF 2124-2205.
RX	MEDLINE=96144829; PubMed=8568869;
RA	Knott V., Downing A.K., Cardy C.M., Handford P.A.;
RT	"Calcium binding properties of an epidermal growth factor-like domain
RT	pair from human fibrillin-1.";
RL	J. Mol. Biol. 255:22-27(1996).
RP	STRUCTURE BY NMR OF 2124-2205.
RX	MEDLINE=96222301; PubMed=8653794;
RA	Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,
RA	Handford P.A.;

RT "Solution structure of a pair of calcium-binding epidermal growth
RT factor-like domains: implications for the Marfan syndrome and other
RT genetic disorders.";
RL Cell 85:597-605(1996).
RN [9]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=96174615; PubMed=8594563;
RA Colloid G., Beroud C., Soussi T., Junien C., Boileau C.;
RT "Software and database for the analysis of mutations in the human
RT FBNI gene.";
RL Nucleic Acids Res. 24:137-141(1996).
RN [10]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=97169383; PubMed=9016526;
RA Colloid-Beroud G., Beroud C., Ales L., Black C., Boxer M., Brock D.J.,
RA Godfrey M., Hayward C., Karttunen C., Mileticz D., Peltonen L.,
RA Richards R.I., Wang W., Juntunen C., Boileau C.;
RT "Marfan Database (second edition): software and database for the
RT analysis of mutations in the human FBNI gene.";
RL Nucleic Acids Res. 25:147-150(1997).
RN [11]
RP REVIEW ON VARIANTS.
RX MEDLINE=98062175; PubMed=9401003;
RA Hayward C., Brock D.J.H.;
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1
RT fibrillogenopathies.";
RL Hum. Mutat. 10:415-423(1997).
RN [12]
RP VARIANT MFS PRO-1137.
RX MEDLINE=91304569; PubMed=1852208;
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,
RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,
RA Curristin S.M., Stetten G., Meyers D.A., Franccomano C.A.;
RT "Marfan syndrome caused by a recurrent de novo missense mutation in
RL the fibrillin gene.";
RL Nature 352:337-339(1991).
RN [13]
RP VARIANTS MFS SER-1249; ARC=1663; SER-2221 AND SER-2307.
RX MEDLINE=93250834; PubMed=1301946;
RA Dietz H.C., Saravia J.M., Pyeritz R.E., Cutting G.R., Franccomano C.A.;
RT "Clustering of fibrillin (FBNI) missense mutations in Marfan syndrome
RT patients at cysteine residues in EGF-like domains.";
RL Hum. Mutat. 1:366-374(1992).
RN [14]
RP VARIANT MFS SER-2307.
RX MEDLINE=92235290; PubMed=1569206;
RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J., Jr.,
RA Corson G.M., Maslen C.L., Sakai L.Y., Franccomano C.A., Cutting G.R.;
RT "Marfan phenotype variability in a family segregating a missense
RT mutation in the epidermal growth factor-like motif of the fibrillin
RT gene.";
RL J. Clin. Invest. 89:1674-1680(1992).
RN [15]
RP VARIANTS MFS ILE-548 AND ALA-723.
RX MEDLINE=94010946; PubMed=8406497;
RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chaberg S.C.,
RA Pyeritz R.E., Franccomano C.A.;
RT "Four novel FBNI mutations: significance for mutant transcript level
RT and EGF-like domain calcium binding in the pathogenesis of Marfan
RT syndrome.";
RL Genomics 17:468-475(1993).
RN [16]
RP VARIANTS MFS SER-2144.
RX MEDLINE=93278402; PubMed=8504310;
RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;
RT "A novel fibrillin mutation in the Marfan syndrome which could
RT disrupt calcium binding of the epidermal growth factor-like module.";
RL Hum. Mol. Genet. 2:475-477(1993).
RN [17]
RP VARIANTS MFS R-862; Y-1117; P-1137 AND F-1589, AND VARIANT A-1148.
RX MEDLINE=94108431; PubMed=8281141;
RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,
RA Berg M.A., Miller D.C., Francke U.;

RT "Mutation screening of complete fibrillin-1 coding sequence: report
RX of five new mutations, including two in 8-cysteine domains.";
RL Hum. Mol. Genet. 2:1813-1821(1993).
RN [18]
RP VARIANTS MFS GLY-217 AMD ARG-2627.
RX MEDLINE=95067970; PubMed=7977366;
RA Karttunen L., Ragnhath M., Leongqvist L., Peltonen L.;
RT "A compound heterozygous Marfan patient: two defective fibrillin
RL alleles result in a lethal phenotype.";
RN Am. J. Hum. Genet. 55:1083-1091(1994).
RP [19]
RX VARIANT EL LYS-2447.
RL MEDLINE=94245249; PubMed=8188302;
RA Leongqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,
RT Peltonen L.;
RL "A novel mutation of the fibrillin gene causing ectopia lentis.";
RN Genomics 19:573-576(1994).
RP [20]
RX VARIANT MFS CYS-627.
RL MEDLINE=94272487; PubMed=8004112;
RA Hayward C., Rae A.L., Porteous M.E.M., Logie I.J., Brock L.J.;
RT "Two novel mutations and a neutral polymorphism in EGF-like domains
of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan
syndrome patients";
RN Hum. Mol. Genet. 3:373-375(1994).
RP [21]
RX VARIANT MFS CYS-122.
RL MEDLINE=94314977; PubMed=8040326;
RA Stehl-Hallegrren C., UKkonen T., Kainulainen K., Kristoferson U.,
RA Saxne T., Tornqvist K., Peltonen L.;
RT "An extra cysteine in one of the non-calcium-binding epidermal growth
factor-like motifs of the FBN1 polypeptide is connected to a novel
variant of Marfan syndrome.";
RN J. Clin. Invest. 94:709-713(1994).
RP [22]
RX VARIANT MFS TYR-1223.
RL MEDLINE=94351682; PubMed=8071963;
RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;
RT "A new missense mutation of fibrillin in a patient with Marfan
syndrome.";
RN J. Med. Genet. 31:338-339(1994).
RP [23]
RX VARIANT MFS HIS-1170.
RL MEDLINE=95174777; PubMed=7870075;
RA Hayward C., Porteous M.E.M., Brock D.J.H.;
RT "A novel mutation in the fibrillin gene (FBN1) in familial
arachnoidectyly";
RN Mol. Cell. Probes 8:325-327(1994).
RP [24]
RX VARIANTS MFS G217;N1023;R1074;Y1242;R1513;E2127;W2151;K2447 AND R2511.
RL MEDLINE=94184366; PubMed=8136837;
RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;
RT "Mutations in the fibrillin gene responsible for dominant ectopia
lentis and neonatal Marfan syndrome.";
RN Nat. Genet. 6:64-69(1994).
RP [25]
RX VARIANT SER-1127.
RL MEDLINE=95282774; PubMed=7762551;

Query Match	12.4%;	Score 363;	DB 1;	Length 2871;
Best Local Similarity	30.8%;	Pred. No. 2e-16;		
Matches 111;	Conservative	22;	Mismatches 105;	Indels 122;
			Gaps	20;
QY	244	DCSVEGGCGEHACNALPGAPRCQCPAGALADRGCTASATQSCNDLCHEHCVPNP---	300	
DB	1200	ECSIMNGGCEFTCTNSEGSGYECSCQGFALMPDRQCT-----DIDE--CEDNPNIC	1249	
QY	301	-----DQGSYSCKMETGYRLAADQREDVDLCLPSPC-PQKCVYTOGGFEDHCYR	353	
DB	1250	DGGQCTNINIGEYRCLCYDDGFMASEDKRTCVDNVEDCLNPNICLSGTCENTKSFICHCIM	1309	
QY	354	NDYLVJGE--CYEPVDPCC--FRANCEYQOQPLNQI--SYLCAVCAEGR-----PIIP	398	

```
Db      1310 GYSGKKGKGTCTD-INECETGAHNCGHNAVCTNTAGSFKSCSPBGWIGDIKCTDLDECS 1368
Qy      399 HEPHRCPMFCNCRACPADCDPNTQAS--CCPEGYILIDGFCTIDIDE-EN----- 447
Db      1369 NGRHMGSQH-----ADC-KNTMGSYRCLCKRGY-TGDSEFTCTDLDCESENLTLCNGCQ 1419
Qy      448 -----GGF-----CS-----GVCHNLPTGEICGPDSALV 473
Db      1420 CLINAPGTYRCEDMGFPADGSKACEDIDECSLPNICVFCTCHNLPLGFLFCECEIGELD 1479
Qy      474 RHFG-----TDCDSGKYVDGSDSGSEPSPSTPGS-----TLPPAVGLV 513
Db      1480 RSGGNCTDYNECLDPITTCISGNCVN-----TPGSYICDCPPPFELNPTRVGCV 1527

RESULT 10
FNBL_MOUSE
ID      FNBL_MOUSE          STANDARD;          PRT; 2871 AA.
AC      061554; 060826;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      FIBRILLIN 1 PRECURSOR.
CN      FNBL OR FRB-1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ox      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95130561; PubMed=7829516;
RA      Yin W., Germiller J., Sanguinetti C., Smiley E., Panglilian T.,
RA      Pereira L., Ramirez F., Bonadio J.;
RT      "Primary structure and developmental expression of Fbn-1, the mouse
RT      fibrillin gene.";
RN      J. Biol. Chem. 270:1798-1806(1995).
[2]
RC      SEQUENCE FROM N.A.
RP      STRAIN=CD-1; TISSUE=Kidney;
RA      Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RA      Submitted (APR-1995) to the EMBL/Genbank/DDBJ databases.
CC      -
CC      -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFILLS
CC      THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFILBRILLS PROVIDE
CC      LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC      -
CC      -1- PM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC      FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC      MICROFILBRILLS (BY SIMILARITY).
CC      -
CC      -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC      EGF-LIKE DOMAINS.
CC      -
CC      -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC      -
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and that statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -
DR      EMBL, L29454; AAA56840.1; -.
DR      EMBL, U22493; AAA64217.1; -.
DR      HSSP; P35555; IAPJ.
DR      MGd; MGI:95489; Fbn1.
DR      InterPro: IPRO00152; Asx_hydroxyl.
DR      InterPro: IPRO00561; EGF-like.
DR      InterPro: IPRO01881; EGF_Ca.
DR      InterPro: IPRO02212; TB.
DR      Pfam; PF00008; EGF; 46.
DR      Pfam; PF00683; TB; 9.
DR      SMART; SM00179; EGF_CA_42.
DR      SMART; SM00001; EGF_like; 4.
DR      PROSITE; PS00010; ASX_HYDROXYL; 43.
DR      PROSITE; PS00022; EGF_1; 2.
```

[illegible]

FT	DISULFID	292	304	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.
FT	DISULFID	315	328	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.
FT	DISULFID	499	513	BY SIMILARITY.
FT	DISULFID	515	528	BY SIMILARITY.
FT	DISULFID	534	546	BY SIMILARITY.
FT	DISULFID	541	555	BY SIMILARITY.
FT	DISULFID	557	570	BY SIMILARITY.
FT	DISULFID	576	587	BY SIMILARITY.
FT	DISULFID	582	596	BY SIMILARITY.
FT	DISULFID	598	611	BY SIMILARITY.
FT	DISULFID	617	628	BY SIMILARITY.
FT	DISULFID	623	637	BY SIMILARITY.
FT	DISULFID	639	652	BY SIMILARITY.
FT	DISULFID	727	739	BY SIMILARITY.
FT	DISULFID	734	748	BY SIMILARITY.
FT	DISULFID	750	763	BY SIMILARITY.
FT	DISULFID	769	781	BY SIMILARITY.
FT	DISULFID	776	790	BY SIMILARITY.
FT	DISULFID	792	805	BY SIMILARITY.
FT	DISULFID	811	821	BY SIMILARITY.
FT	DISULFID	816	830	BY SIMILARITY.
FT	DISULFID	832	845	BY SIMILARITY.
FT	DISULFID	914	926	BY SIMILARITY.
FT	DISULFID	921	935	BY SIMILARITY.
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1111	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.
FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	1658	1672	BY SIMILARITY.

Query Match 12.48; Score 361; DB 1; Length 2871;

Best Local Similarity 29.48; Pred. No. 2.7e-16;
Matches 117; Conservative 26; Mismatches 127;
Indels 128; Gaps 21.

Oy	212	SAAVAPRLGL-----OLMTCAPRGAAVQGHMARBARPCAMDCSVENGGCEHACNAIVGAPRC	265
Dd	1162	SANLCSPHGRCVNLIGTKYQCACNGFYHPTRDLRFLCVUIDECSTINNGCETFFCTINSOGSYEC	1222
Oy	266	QCPAGALQADGRSCTASATQSCNDLCEHFVCVPNP-----DQPSYSQMCETGYRL	316
Dd	1222	SCQGFGLMVDQRSCST-----DIDQ--CEDNPNPICDGQOCTNIPGEYRCLCYDFEWA	1271
Oy	317	AARDHREDVEDDCLLESPPC-PQRCVNTOGGFCHCYRPNYDLVDGE--CVERVDDCC-FR	371
Dd	1272	SEDKTTCVVDNECDLANPNTICLSGTCENTKSGSFICHDMGSGKKGTGCTD-INDEIGA	1330
Oy	372	ANCEYCOCPLNQT-SYLCAEAEGFA-----PIPEPHRCOMFCNOTACPADCDPN	420
Dd	1331	HNCRRHAVCTNTPAGSFRCSSCPWIGDIGIKCTDLDECNSGTHMSQH-----ADC-KN	1387
Oy	421	TQAAS--CECGEPGLYLDGFCITPIDEC-EN-----GGR-----	450
Dd	1383	TMGSRCLCKDXG-TGDGFTCTDLDESENILNCGMGCLNABGRCIREECDMGFPVSADG	1444
Oy	451	-----CS-----GVCHNLPGTFPECICGPDSALVRHIIG-----TDGSGK	484
Dd	1442	KACEDIDECSLPNVICVGCTCHNLPGFLFRCEIGEYELDRSGGNCVTDNEDLPJTCTISGN	1503
Oy	485	VDDGSDSGSEPPTPGS-----TLTPRVAGLY	513
Dd	1502	CVN-----TPGSYTCDSCSPDEFELNPRVGCV	1527
<hr/>			
RESULT	11		
ID	MATR2_HUMAN	STANDARD;	PRT; 956 AA.
AC	000339; Q9NSZ1;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	20-AUG-2001 (Rel. 40, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	MATRILIN-2 PRECURSOR.		
GN	Matn2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Eukaryota; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RN	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).		
RA	Pubmed-11124542;		
RA	Muratoglu S., Krysan K., Balazs M., Sheng H., Zakany R., Modis L.,		
RA	Kiss I., Deak F.:		
RT	"Primary structure of human matrilin-2, chromosome location of the		
RT	MATN2 gene and conservation of an AT-rich Intron in matrilin genes.";		
RL	Cytogenet. Cell Genet. 90:323-327(2000).		
RN	[2]		
RN	SEQUENCE OF 644-935 FROM N.A.		
RX	MEDLINE=97238863; Pubmed=9083061;		
RA	Deak F., Piecha D., Bachrati C., Paulsson M., Kiss I.:		
RT	"Primary structure and expression of matrilin-2, the closest relative		
RT	of cartilage matrix protein within the von Willebrand factor type A-		
RL	J. Biol. Chem. 272:9268-9274(1997).		
RN	[3]		
RP	SEQUENCE OF 244-956 FROM N.A.		
RC	TISSUE=Testis		
RA	Dusterhoeft A., Tauber J., Meves H.-W., Gassenhuber J., Wiemann S.:		
CC	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: SECRETED.		
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A		
CC	SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC	-1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 2 WFPA DOMAINS.		

[illegible]

DR	PROSITE; PS50086; ANK_REPEAT; 5.
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR	PROSITE; PS00010; ASX_HYDROXYL; 11.
DR	PROSITE; PS00022; EGF_1; 28.
DR	PROSITE; PS01186; EGF_2; 21.
DR	PROSITE; PS01187; EGF_CA; 9.
KW	differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane; Glycoprotein; Proto-oncogene; ANK Repeat; Signal.
FT	CHAIN 1 20
FT	DOMAIN 1 1964
FT	DOMAIN 21 1443
FT	TRANSMEM 1444 1464
FT	DOMAIN 1465 1964
FT	DOMAIN 21 60
FT	DOMAIN 61 112
FT	DOMAIN 115 152
FT	DOMAIN 153 189
FT	DOMAIN 211 229
FT	DOMAIN 231 271
FT	DOMAIN 309 350
FT	DOMAIN 311 350
FT	DOMAIN 352 388
FT	DOMAIN 389 427
FT	DOMAIN 429 470
FT	DOMAIN 472 508
FT	DOMAIN 510 546
FT	DOMAIN 548 584
FT	DOMAIN 586 622
FT	DOMAIN 622 656
FT	DOMAIN 658 686
FT	DOMAIN 688 724
FT	DOMAIN 724 762
FT	DOMAIN 764 800
FT	DOMAIN 803 839
FT	DOMAIN 841 877
FT	DOMAIN 878 924
FT	DOMAIN 926 962
FT	DOMAIN 964 1000
FT	DOMAIN 1002 1040
FT	DOMAIN 1042 1081
FT	DOMAIN 1083 1122
FT	DOMAIN 1126 1167
FT	REPEAT 1168 1208
FT	REPEAT 1209 1242
FT	REPEAT 1243 1282
FT	REPEAT 1283 1657
FT	REPEAT 1658 1661
FT	REPEAT 1661 1724
FT	REPEAT 1695 1724
FT	REPEAT 1728 1757
FT	REPEAT 1761 1790
FT	DISULFID 25 38
FT	DISULFID 32 48
FT	DISULFID 50 59
FT	DISULFID 65 77
FT	DISULFID 71 100
FT	DISULFID 102 111
FT	DISULFID 113 130
FT	DISULFID 124 140
FT	DISULFID 142 151
FT	DISULFID 157 168
FT	DISULFID 162 177
FT	DISULFID 179 188
FT	DISULFID 195 218
FT	DISULFID 202 217
FT	DISULFID 219 228
FT	DISULFID 235 246
FT	DISULFID 240 259
FT	DISULFID 261 270
FT	DISULFID 277 288
FT	DISULFID 282 297
FT	DISULFID 299 308
FT	DISULFID 315 329
FT	DISULFID 323 338
FT	DISULFID 338 349
FT	DISULFID 349 360
FT	DISULFID 360 371
FT	DISULFID 371 382
FT	DISULFID 382 393
FT	DISULFID 393 404
FT	DISULFID 404 415
FT	DISULFID 415 426
FT	DISULFID 426 437
FT	DISULFID 437 448
FT	DISULFID 448 459
FT	DISULFID 459 470
FT	DISULFID 470 481
FT	DISULFID 481 492
FT	DISULFID 492 503
FT	DISULFID 503 514
FT	DISULFID 514 525
FT	DISULFID 525 536
FT	DISULFID 536 547
FT	DISULFID 547 558
FT	DISULFID 558 569
FT	DISULFID 569 580
FT	DISULFID 580 591
FT	DISULFID 591 602
FT	DISULFID 602 613
FT	DISULFID 613 624
FT	DISULFID 624 635
FT	DISULFID 635 646
FT	DISULFID 646 657
FT	DISULFID 657 668
FT	DISULFID 668 679
FT	DISULFID 679 690
FT	DISULFID 690 701
FT	DISULFID 701 712
FT	DISULFID 712 723
FT	DISULFID 723 734
FT	DISULFID 734 745
FT	DISULFID 745 756
FT	DISULFID 756 767
FT	DISULFID 767 778
FT	DISULFID 778 789
FT	DISULFID 789 800
FT	DISULFID 800 811
FT	DISULFID 811 822
FT	DISULFID 822 833
FT	DISULFID 833 844
FT	DISULFID 844 855
FT	DISULFID 855 866
FT	DISULFID 866 877
FT	DISULFID 877 888
FT	DISULFID 888 899
FT	DISULFID 899 910
FT	DISULFID 910 921
FT	DISULFID 921 932
FT	DISULFID 932 943
FT	DISULFID 943 954
FT	DISULFID 954 965
FT	DISULFID 965 976
FT	DISULFID 976 987
FT	DISULFID 987 998
FT	DISULFID 998 1009
FT	DISULFID 1009 1020
FT	DISULFID 1020 1031
FT	DISULFID 1031 1042
FT	DISULFID 1042 1053
FT	DISULFID 1053 1064
FT	DISULFID 1064 1075
FT	DISULFID 1075 1086
FT	DISULFID 1086 1097
FT	DISULFID 1097 1108
FT	DISULFID 1108 1119
FT	DISULFID 1119 1130
FT	DISULFID 1130 1141
FT	DISULFID 1141 1152


```

FT DISULFID 340 349 BY SIMILARITY.
FT DISULFID 356 376 BY SIMILARITY.
FT DISULFID 361 376 BY SIMILARITY.
FT DISULFID 378 387 BY SIMILARITY.
FT DISULFID 393 404 BY SIMILARITY.
FT DISULFID 398 415 BY SIMILARITY.
FT DISULFID 417 426 BY SIMILARITY.
FT DISULFID 433 449 BY SIMILARITY.
FT DISULFID 443 458 BY SIMILARITY.
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
FT DISULFID 481 496 BY SIMILARITY.
FT DISULFID 498 507 BY SIMILARITY.
FT DISULFID 514 525 BY SIMILARITY.
FT DISULFID 519 534 BY SIMILARITY.
FT DISULFID 536 545 BY SIMILARITY.
FT DISULFID 552 563 BY SIMILARITY.
FT DISULFID 557 572 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
FT DISULFID 590 601 BY SIMILARITY.
FT DISULFID 595 610 BY SIMILARITY.
FT DISULFID 612 621 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 655 BY SIMILARITY.
FT DISULFID 662 669 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 676 685 BY SIMILARITY.
FT DISULFID 692 703 BY SIMILARITY.
FT DISULFID 712 723 BY SIMILARITY.
FT DISULFID 714 723 BY SIMILARITY.
FT DISULFID 730 741 BY SIMILARITY.
FT DISULFID 735 750 BY SIMILARITY.
FT DISULFID 752 761 BY SIMILARITY.
FT DISULFID 768 779 BY SIMILARITY.
FT DISULFID 773 788 BY SIMILARITY.
FT DISULFID 790 799 BY SIMILARITY.
FT DISULFID 807 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 829 838 BY SIMILARITY.
FT DISULFID 845 856 BY SIMILARITY.
FT DISULFID 860 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 882 903 BY SIMILARITY.
FT DISULFID 897 912 BY SIMILARITY.
FT DISULFID 914 923 BY SIMILARITY.
FT DISULFID 930 941 BY SIMILARITY.
FT DISULFID 935 950 BY SIMILARITY.
FT DISULFID 952 961 BY SIMILARITY.
FT DISULFID 968 979 BY SIMILARITY.
FT DISULFID 973 988 BY SIMILARITY.
FT DISULFID 990 999 BY SIMILARITY.
FT DISULFID 1006 1019 BY SIMILARITY.
FT DISULFID 1011 1028 BY SIMILARITY.
FT DISULFID 1030 1039 BY SIMILARITY.
FT DISULFID 1046 1057 BY SIMILARITY.
FT DISULFID 1051 1069 BY SIMILARITY.
FT DISULFID 1071 1080 BY SIMILARITY.
FT DISULFID 1087 1098 BY SIMILARITY.

```

Query Match 11.3%; Score 329.5; DB 1; Length 1964;
 Best Local Similarity 26.3%; Pred. No. 2.2e-14;
 Matches 128; Conservative 33; Mismatches 143; Indels 183; Gaps 31;

```

OY 134 GYLCAVSAAEATVSEPTMEQOCEVADFLCE-FHPATCRLANVEPGAAAAVSIT 192
DB 35 GGTCLSLSGGIC-----QC---APGLVTCGFPDPCR----- 66
OY 193 YGTPFAARGADFOAL---FVGSSAAVAPLGLMCTAPG---AVQGHMAREAPGAMDC 245
DB 67 -DTOLCKNGSGQALLPFPSSRSPPLTPHFSCTCPSGFTGDCQTHLEICPPSF-C 124
OY 246 SYVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVPNPDQF- 304

```

```

DB 125 S--NGG--HCYVQASGRPOCCCEPGWT-----GEQO-----LRDFCSANCANMG 166
OY 305 -----SYSCMETGYRLAADQHRCE-DVDDCIIEPSPQ--RCVNTGGFEGHC--- 351
DB 167 VCLATVPQICRCRPPGF-----EGHTCENDIINECFLEPGRCPOGTSCHMTLSGYQCLCPVG 222
OY 352 --YPTNDLVNDEGEVEVDPCFRANCERYOCPL---NOTSYLCVACAGFAPRIHE----- 400
DB 223 QEGPOCKLRKGAC--PGSCLNG--TCOLVPEGHSTFHLCPCPGFGLDCEMNPDC 277
OY 401 -PHRCOMFCQGTACPADCPNTOASCEPGYLDGFCIT--DIDCE-----NGGF 450
DB 278 VRHQCO---NGATCLDGLDITY--CLCRXTM---KGMDCSEDIDCECARGPPRCRNGST 328
OY 451 C-----SGVCHNLPGEFCICG----- 468
DB 329 CQMTAGSFHCVCVSGMGACCEENLDCAATCAPGSTCIDRVGSFSCLPGRGTLCH 388
OY 469 --DSALVR--HITDSDSGKVDG-----GDSG-----GPPSP----- 499
DB 389 LEDMCLSQPCRVNAQCSFTPLTGSTYLICQPGYSGSTCHQDIDCECOMAQGSPCEHGS 448
OY 500 ---TPGS 503
DB 449 CINTPGS 455

RESULT 14
MTN2_MOUSE STANDARD: PRT: 956 AA.
AC 008746;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MATRILIN-2 PRECURSOR.
GN MATN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C; TISSUE=Limb;
RX MEDLINE=97238663; PubMed=9083061;
RA Deak F., Plecha D., Bachratl C., Paulsson M., Kiss I.;
RT "Primary structure and expression of matrilin-2, the closest relative
RT of cartilage matrix protein within the von Willebrand factor type A-
RT like module superfamily.";
RL J. Biol. Chem. 272:9268-9274(1997).
CC -I- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
CC CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
CC OSTEOBLAST CELL LINES.
CC -I- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 VMFA DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DB EMBL; U69262; AAC53163.1; -
DB HSSP; P07204; 1ZAQ.
DB MGD; MGI:109613; Matn2.
DB InterPro; IPR000152; Asx_hydroxyl.
DB InterPro; IPR000561; EGF-like.
DB InterPro; IPR001881; EGF-Ca.
DB InterPro; IPR002035; VMFA.

```


FT	SITE	734	736	BINDING PROTEIN 1.
FT	SITE	1575	1577	CLEAVAGE (POTENTIAL).
FT	PROPEP	1578	1712	CLEAVAGE (POTENTIAL).
FT	DOMAIN	181	213	POTENTIAL.
FT	DOMAIN	391	423	EGF-LIKE 1.
FT	REPEAT	551	604	INTERNAL REPEAT 1.
FT	DOMAIN	618	658	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT	REPEAT	671	721	INTERNAL REPEAT 2.
FT	DOMAIN	865	906	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	907	948	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	949	989	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	990	1029	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1030	1070	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1071	1111	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1112	1151	EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1153	1193	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1194	1235	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1236	1277	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1278	1320	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT	REPEAT	1340	1392	INTERNAL REPEAT 3.
FT	DOMAIN	1415	1457	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1458	1498	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT	REPEAT	1517	1568	INTERNAL REPEAT 4.
FT	DOMAIN	1612	1652	EGF-LIKE 17.
FT	DOMAIN	1653	1697	EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT	DISULFID	185	195	BY SIMILARITY.
FT	DISULFID	189	201	BY SIMILARITY.
FT	DISULFID	203	212	BY SIMILARITY.
FT	DISULFID	395	405	BY SIMILARITY.
FT	DISULFID	399	411	BY SIMILARITY.
FT	DISULFID	413	422	BY SIMILARITY.
FT	DISULFID	622	633	BY SIMILARITY.
FT	DISULFID	628	642	BY SIMILARITY.
FT	DISULFID	644	657	BY SIMILARITY.
FT	DISULFID	869	881	BY SIMILARITY.
FT	DISULFID	876	890	BY SIMILARITY.
FT	DISULFID	892	905	BY SIMILARITY.
FT	DISULFID	911	923	BY SIMILARITY.
FT	DISULFID	918	932	BY SIMILARITY.
FT	DISULFID	934	947	BY SIMILARITY.
FT	DISULFID	953	964	BY SIMILARITY.
FT	DISULFID	959	973	BY SIMILARITY.
FT	DISULFID	976	988	BY SIMILARITY.
FT	DISULFID	994	1005	BY SIMILARITY.
FT	DISULFID	1000	1014	BY SIMILARITY.
FT	DISULFID	1017	1028	BY SIMILARITY.
FT	DISULFID	1034	1045	BY SIMILARITY.
FT	DISULFID	1040	1054	BY SIMILARITY.
FT	DISULFID	1056	1069	BY SIMILARITY.
FT	DISULFID	1075	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1110	BY SIMILARITY.
FT	DISULFID	1116	1127	BY SIMILARITY.
FT	DISULFID	1122	1136	BY SIMILARITY.
FT	DISULFID	1138	1151	BY SIMILARITY.
FT	DISULFID	1157	1169	BY SIMILARITY.
FT	DISULFID	1164	1178	BY SIMILARITY.
FT	DISULFID	1180	1192	BY SIMILARITY.
FT	DISULFID	1198	1210	BY SIMILARITY.
FT	DISULFID	1204	1219	BY SIMILARITY.
FT	DISULFID	1221	1234	BY SIMILARITY.
FT	DISULFID	1240	1252	BY SIMILARITY.
FT	DISULFID	1246	1261	BY SIMILARITY.
FT	DISULFID	1263	1276	BY SIMILARITY.
FT	DISULFID	1282	1294	BY SIMILARITY.
FT	DISULFID	1289	1303	BY SIMILARITY.
FT	DISULFID	1305	1319	BY SIMILARITY.
FT	DISULFID	1419	1432	BY SIMILARITY.
FT	DISULFID	1427	1441	BY SIMILARITY.
FT	DISULFID	1443	1456	BY SIMILARITY.
FT	DISULFID	1462	1473	BY SIMILARITY.
FT	DISULFID	1468	1482	BY SIMILARITY.

FT	DISULFID	1484	1497	BY SIMILARITY.
FT	DISULFID	1616	1627	BY SIMILARITY.
FT	DISULFID	1622	1636	BY SIMILARITY.
FT	DISULFID	1638	1651	BY SIMILARITY.
FT	DISULFID	1657	1672	BY SIMILARITY.
FT	DISULFID	1667	1681	BY SIMILARITY.
FT	DISULFID	1683	1696	BY SIMILARITY.
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	370	370	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	416	416	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1042	1042	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1242	1242	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1357	1357	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	1712 AA:	186598 MW:	650BCEA691FDD34 CRC64;

Query Match 11.2%; Score 327.5; DB 1; Length 1712;
Best Local Similarity 28.8%; Pred. No. 2.7e-14;
Matches 106; Conservative 44; Mismatches 133; Indels 85; Gaps 25;

QY	177	PLAVPFGAAAAAVSYTGTTPFAKADQALPVGS----	SAAVAPLGLQLM--CLAPGCA	230	
DB	831	PPVETETPPVVEVA-----	PEGSTASASOVIATPYTETNEICVNNDI	875	
QY	231	VOGHAREAPGAMDSVENG-----	GCEHA-----	CNAIPGARPCOP	268
DB	876	CGAGHCINLPVAYTICTEGYKFSHQQRKCIDIDBCAQAHLCSGKRENTGSLTICP		935	
QY	269	AGALQADGRSGTASATGSCNDLC--EHRCPNPDPGYSYC--MCETGYRLAADQHRCEVD		326	
DB	936	AGFIASEGSGNCIDVDECLRPDYCRDGRCI--	NTAGAFRCPEYCDSGYRMSRGRH-CEDI	991	
QY	327	DDCILEPSPCP--QRVUNPQSGFEC--HCYPNVDLVGCEVPEPDPREFRAN--	CEYQCGOLN	382	
DB	992	DEC-LTPETCEPEQCVNPSGSTQCPCTGEGFGMNGCCID--	VDECLQPRVCTNGSCTNL--	1048	
QY	383	QTSYLCAVEAGFAPRPHNP-----RCOM--	FCNQTAPRADCDPMTQAS--	CEPEGYIL	433
DB	1049	ESYWSGCHKGYSPPRRHRCDDIDECQGNLCNMGGC-----	KNTDGSFRCTGGQGYQL	1103	
QY	434	----DGFICTIDIDCEKMGFGCS--GVCNNLPGTEFGICGPD--	SALVRHNG--	ITDC--D	481
DB	1104	SAKQDQ-----CEDIDICEHRHLCSHGQCRNTGEGFQCLCNOGYRASVLDGDHEDINECLED		1160	
QY	482	SGKYDGGD	489		
DB	1161	SSVCQGGD	1168		

Search completed: May 8, 2002, 12:45:14
Job time: 350 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 8, 2002, 12:37:49 ; Search time 206.9 Seconds

(without alignments)
876.013 Million cell updates/sec

Title: US-09-509-994-1

Sequence: 1 MLGVLVGLALAGLGPAP.....PSPPGSLTPPAVLVHSG 516

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US060_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US061_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US062_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US063_COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US064_COMB.pep:*
29: /cgn2_6/ptodata/2/paa/US065_COMB.pep:*
30: /cgn2_6/ptodata/2/paa/US066_COMB.pep:*
31: /cgn2_6/ptodata/2/paa/US067_COMB.pep:*
32: /cgn2_6/ptodata/2/paa/US068_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	516	27	US-09-509-994-1

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
2916	100.0	575	22	US-07-689-936A-7	Sequence 7, Appl1																																						
2916	100.0	575	22	US-09-880-464-2	Sequence 2, Appl1																																						
2916	100.0	575	23	US-09-938-405-2	Sequence 2, Appl1																																						
2912	99.9	575	27	US-09-509-994-2	Sequence 2, Appl1																																						
2912	99.9	575	3	US-07-730-975-2	Sequence 2, Appl1																																						
2912	99.9	575	3	US-07-948-759-4	Sequence 2, Appl1																																						
2912	99.9	575	23	US-09-948-933-228	Sequence 228, App																																						
2912	99.9	682	17	US-09-948-933-318	Sequence 318, App																																						
2824	96.8	497	17	US-09-331-793-4	Sequence 4, Appl1																																						
2770	95.0	494	3	US-07-796-336A-14	Sequence 14, Appl1																																						
2768	94.9	494	3	US-07-796-336A-16	Sequence 16, Appl1																																						
2746	94.2	486	6	US-08-212-384-4	Sequence 4, Appl1																																						
2690	92.2	475	3	US-07-835-436C-1	Sequence 1, Appl1																																						
2687	92.1	475	3	US-07-835-436-1	Sequence 1, Appl1																																						
2687	92.1	475	3	US-07-835-436A-1	Sequence 1, Appl1																																						
2686	92.1	475	3	US-07-835-436-2	Sequence 2, Appl1																																						
2686	92.1	475	3	US-07-835-436A-2	Sequence 2, Appl1																																						
2686	92.1	475	3	US-07-835-436C-2	Sequence 2, Appl1																																						
2684	92.0	476	3	US-07-796-336A-2	Sequence 2, Appl1																																						
2682	92.0	476	3	US-07-796-336A-17	Sequence 17, Appl1																																						
2665	91.4	468	6	US-08-212-384-1	Sequence 1, Appl1																																						
2630	90.2	462	3	US-07-897-194-1	Sequence 1, Appl1																																						
2630	90.2	462	3	US-07-897-194A-1	Sequence 1, Appl1																																						
2619	89.8	462	3	US-07-897-194-3	Sequence 3, Appl1																																						
2615	89.7	462	3	US-07-897-194-4	Sequence 4, Appl1																																						
2615	89.7	462	3	US-07-897-194-5	Sequence 5, Appl1																																						
2598	89.1	462	3	US-07-897-194-6	Sequence 6, Appl1																																						
2592	88.9	456	3	US-07-835-436C-3	Sequence 3, Appl1																																						
2592	88.9	456	3	US-07-835-436C-3	Sequence 3, Appl1																																						
2592	88.9	456	3	US-07-835-436C-3	Sequence 3, Appl1																																						
2592	88.9	456	3	US-07-835-436C-4	Sequence 4, Appl1																																						
2592	88.9	456	3	US-07-835-436C-4	Sequence 4, Appl1																																						
2591	88.9	462	3	US-07-897-194-10	Sequence 10, Appl1																																						
2576	88.3	462	3	US-07-897-194-2	Sequence 2, Appl1																																						
2571	88.2	462	3	US-07-897-194-16	Sequence 16, Appl1																																						
2571	88.2	462	3	US-07-897-194-18	Sequence 18, Appl1																																						
2566	88.0	462	3	US-07-897-194-14	Sequence 14, Appl1																																						
2561	87.8	462	3	US-07-897-194-13	Sequence 13, Appl1																																						
2558	87.7	462	3	US-07-897-194-12	Sequence 12, Appl1																																						
2543	87.2	446	3	US-07-835-436-5	Sequence 5, Appl1																																						
2543	87.2	446	3	US-07-835-436A-5	Sequence 5, Appl1																																						

ALIGNMENTS

RESULT 1

US-09-509-994-1

Sequence 1, Application US/09509994

GENERAL INFORMATION:

APPLICANT: YOKOZAWA, AKIRA

APPLICANT: YOI, MASAKI

APPLICANT: MURATA, TOMOYO

APPLICANT: TSURUTA, KAZUHISA

APPLICANT: SHIMIZU, HIROMOMO

TITLE OF INVENTION: METHOD FOR KEEPING THE QUALITY OF AQUEOUS PARENTERAL SOLUTION OF THROMBOMODULIN IN STORAGE AND DISTRIBUTION

FILE REFERENCE: KP-8753

CURRENT APPLICATION NUMBER: US/09/509,994

PRIOR FILING DATE: 2000-05-08

PRIOR APPLICATION NUMBER: PCT/JP98/04609

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: JP 9-281659

PRIOR FILING DATE: 1997-10-15

PRIOR APPLICATION NUMBER: JP 9-308523

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 516

APPLICANTS

```

; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Partial amino acid sequences of a human
; OTHER INFORMATION: thrombomodulin
US-09-509-994-1

```

```

Query Match          100.0%; Score 2916; DB 27; Length 516;
Best Local Similarity 100.0%; Pred. No. 4e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVYVIGALALAGLGFPAAPAPQPGSGQVEHDFALYPPATFLNASQICDGLRGLM 60
DB 1 MGVYVIGALALAGLGFPAAPAPQPGSGQVEHDFALYPPATFLNASQICDGLRGLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
DB 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
QY 181 EPGAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAAPLGLQIMCTAPGAVQGHMAREAP 240
DB 181 EPGAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAAPLGLQIMCTAPGAVQGHMAREAP 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHCYPNTDLYDG 360
DB 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHCYPNTDLYDG 360
QY 361 BCEVPEVDFCFRANCYOCOPLNQSTYLCVAGFAPITHEPHRCOMFQNCAPACDQPN 420
DB 361 BCEVPEVDFCFRANCYOCOPLNQSTYLCVAGFAPITHEPHRCOMFQNCAPACDQPN 420
QY 421 TQASCEPEGYTLDDGFLCTDIDECENGFCGVCVCHNLPGTFECICGPDASLVHIGTDC 480
DB 421 TQASCEPEGYTLDDGFLCTDIDECENGFCGVCVCHNLPGTFECICGPDASLVHIGTDC 480
QY 481 DSGKVDGSDSGSEPPSPFPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPFPGSTLTPPAVGLVHSG 516

```

RESULT 2

```

; Sequence 7, Application US/07689936A
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Gomi, Komakazu
; APPLICANT: Ogawa, Kohel
; TITLE OF INVENTION: An Isolated Physiologically Active
; TITLE OF INVENTION: Human Thrombomodulin Polypeptide
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington Street P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,936A

```

```

; FILING DATE: 19910521
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP90/01234
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-0369
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "thrombomodulin preprotein"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..18
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "signal peptide"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 517..575
; OTHER INFORMATION: /label= domain
; OTHER INFORMATION: /note= "membrane-binding and cytoplasmic domains"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 366..480
; OTHER INFORMATION: /label= domain
; OTHER INFORMATION: /note= "enzyme active site and chondroitin or
; OTHER INFORMATION: chondroitin-sulfate attachment region"
US-07-689-936A-7

```

```

Query Match          100.0%; Score 2916; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.5e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGVYVIGALALAGLGFPAAPAPQPGSGQVEHDFALYPPATFLNASQICDGLRGLM 60
DB 1 MGVYVIGALALAGLGFPAAPAPQPGSGQVEHDFALYPPATFLNASQICDGLRGLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
DB 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
QY 181 EPGAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAAPLGLQIMCTAPGAVQGHMAREAP 240
DB 181 EPGAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAAPLGLQIMCTAPGAVQGHMAREAP 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHCYPNTDLYDG 360
DB 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHCYPNTDLYDG 360

```

QY 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 Db 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 QY 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 Db 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||
 Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||

RESULT 3
 US-09-880-464-2 09880484

; Sequence 2, Application US/09880464
 ; GENERAL INFORMATION:
 ; APPLICANT: Morser, David
 ; APPLICANT: Morser, Michael J
 ; APPLICANT: Nagashima, Mariako
 ; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use
 ; FILE REFERENCE: Improved Thrombomodulin Analogs
 ; CURRENT APPLICATION NUMBER: US/09/880,464
 ; PRIOR FILING DATE: 2001-08-21
 ; PRIOR APPLICATION NUMBER: 60/213,678
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 575
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-464-2

Query Match 100.0%; Score 2916; DB 22; Length 575;
 Best Local Similarity 100.0%; Pred. No. 4.5e-193;
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRHLM 60
 |||||||
 Db 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRHLM 60
 |||||||
 QY 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKLGPLGFQWVTGDNNTSYS 120
 |||||||
 Db 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKLGPLGFQWVTGDNNTSYS 120
 |||||||
 QY 121 RMARLDNGAPLGLPCLVAVSAEAATVPSEPIWEBOOCEVKADGFLCEHFHPTACRPLAV 180
 |||||||
 Db 121 RMARLDNGAPLGLPCLVAVSAEAATVPSEPIWEBOOCEVKADGFLCEHFHPTACRPLAV 180
 |||||||
 QY 181 EPGAAAANVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
 |||||||
 Db 181 EPGAAAANVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
 |||||||
 QY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 Db 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 QY 301 DQPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
 |||||||
 Db 301 DQPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
 |||||||
 QY 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 Db 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 QY 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 Db 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||

Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||

RESULT 4
 US-09-938-405-2
 ; Sequence 2, Application US/09938405
 ; GENERAL INFORMATION:
 ; APPLICANT: Morser, Barry W.
 ; APPLICANT: Morser, Michael J.
 ; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Inju
 ; FILE REFERENCE: 51960AUSM1
 ; CURRENT APPLICATION NUMBER: US/09/938,405
 ; PRIOR FILING DATE: 2001-08-23
 ; PRIOR APPLICATION NUMBER: 60/229,714
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 575
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-938-405-2

Query Match 100.0%; Score 2916; DB 23; Length 575;
 Best Local Similarity 100.0%; Pred. No. 4.5e-193;
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRHLM 60
 |||||||
 Db 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRHLM 60
 |||||||
 QY 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKLGPLGFQWVTGDNNTSYS 120
 |||||||
 Db 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKLGPLGFQWVTGDNNTSYS 120
 |||||||
 QY 121 RMARLDNGAPLGLPCLVAVSAEAATVPSEPIWEBOOCEVKADGFLCEHFHPTACRPLAV 180
 |||||||
 Db 121 RMARLDNGAPLGLPCLVAVSAEAATVPSEPIWEBOOCEVKADGFLCEHFHPTACRPLAV 180
 |||||||
 QY 181 EPGAAAANVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
 |||||||
 Db 181 EPGAAAANVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
 |||||||
 QY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 Db 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 QY 301 DQPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
 |||||||
 Db 301 DQPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
 |||||||
 QY 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 Db 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPRPHRHRCOMFCNQTACPADCDPN 420
 |||||||
 QY 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 Db 421 TQASCECEGTYLDDGFTCTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVHIGTDC 480
 |||||||
 QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||
 Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
 |||||||

RESULT 5
 US-09-509-994-2
 ; Sequence 2, Application US/0950994
 ; GENERAL INFORMATION:
 ; APPLICANT: YUI, MASAKI
 ; APPLICANT: YOKOZAKA, AKIRA
 ; APPLICANT: MURATA, TOMOYO
 ; APPLICANT: TSURUTA, KAZUHISA

APPLICANT: SHIMIZU, HIROKAZU
TITLE OF INVENTION: METHOD FOR KEEPING THE QUALITY OF AQUEOUS PARENTERAL
FILE OF INVENTION: SOLUTION OF THROMBOMODULIN IN STORAGE AND DISTRIBUTION
FILE REFERENCE: KP-8753
CURRENT APPLICATION NUMBER: US/09/509,994
CURRENT FILING DATE: 2000-05-08
PRIOR APPLICATION NUMBER: PCT/JP98/04609
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: JP 9-281659
PRIOR FILING DATE: 1997-10-15
PRIOR APPLICATION NUMBER: JP 9-308523
PRIOR FILING DATE: 1997-11-11
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 516
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Partial amino acid sequences of a human
US-09-509-994-2

Query Match 99.98% Score 2912; DB 27; Length 516;
Best Local Similarity 99.8%; Pred. No. 7.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
DB 1 MGVLVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGPFOWVTGDNNTSYS 120
DB 61 TVRSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGPFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGLCPVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 180
QY 181 EPGAAAASITTYGTPPARAGADFQALPVGSSAAVAPLGIQIMCTAPGAVOGHMAREAP 240
DB 181 EPGAAAASITTYGTPPARAGADFQALPVGSSAAVAPLGIQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCEVNGGCEHACNAITGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCEVNGGCEHACNAITGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQRCVNTQGGFECCHYPNVDLYDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQRCVNTQGGFECCHYPNVDLYDG 360
QY 361 ECEVPDPCFRANCEYQCPQLNQTSTLYCAEGBFAPLPHPHRCQMFQNOTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCPQLNQTSTLYCAEGBFAPLPHPHRCQMFQNOTACPADCDPN 420
QY 421 TQASCEPREGYIIDDGFICIDIDECENGFCGSGVCHNLPTGFCICGSPSALVRIHGTDC 480
DB 421 TQASCEPREGYIIDDGFICIDIDECENGFCGSGVCHNLPTGFCICGSPSALVRIHGTDC 480
QY 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

RESULT 6
US-07-730-975-2
Sequence 2, Application US/07730975
GENERAL INFORMATION:
APPLICANT: MORSE, MICHAEL J.
APPLICANT: GLASER, CHARLES
TITLE OF INVENTION: SOLUBLE ANALOGS OF THROMBOMODULIN
NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:
ADDRESSEE: BERLEX BIOSCIENCES
STREET: 1501 HARBOR BAY PARKWAY
CITY: ALAMEDA
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94501
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/730,975
FILING DATE: 19910729
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: CHING, EDWIN P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: A-0155A US
TELEPHONE: (415) 266-7476
TELEFAX: (415) 266-7400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-730-975-2

Query Match 99.98% Score 2912; DB 3; Length 575;
Best Local Similarity 99.8%; Pred. No. 8.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
DB 1 MGVLVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGPFOWVTGDNNTSYS 120
DB 61 TVRSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGPFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGLCPVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 180
QY 181 EPGAAAASITTYGTPPARAGADFQALPVGSSAAVAPLGIQIMCTAPGAVOGHMAREAP 240
DB 181 EPGAAAASITTYGTPPARAGADFQALPVGSSAAVAPLGIQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCEVNGGCEHACNAITGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCEVNGGCEHACNAITGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQRCVNTQGGFECCHYPNVDLYDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQRCVNTQGGFECCHYPNVDLYDG 360
QY 361 ECEVPDPCFRANCEYQCPQLNQTSTLYCAEGBFAPLPHPHRCQMFQNOTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCPQLNQTSTLYCAEGBFAPLPHPHRCQMFQNOTACPADCDPN 420
QY 421 TQASCEPREGYIIDDGFICIDIDECENGFCGSGVCHNLPTGFCICGSPSALVRIHGTDC 480
DB 421 TQASCEPREGYIIDDGFICIDIDECENGFCGSGVCHNLPTGFCICGSPSALVRIHGTDC 480
QY 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516


```

RESULT 7
US-07-949-759-4
: Sequence 4, Application US/07949759
: GENERAL INFORMATION:
: APPLICANT: Beeler, David L.
: APPLICANT: Jackman, Robert W
: APPLICANT: Rosenberg, Robert B
: APPLICANT: Fritze, Linda
: APPLICANT: Soff, Gerald
: TITLE OF INVENTION: Thrombomodulin of Mammalian Origin
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Milltia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/949,759
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION NUMBER:
: APPLICATION NUMBER: US 07/675,307
: FILING DATE: 26-MAR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,227
: REFERENCE/DOCKET NUMBER: MT-4545A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 617-861-9540
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 575 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-07-949-759-4

```

```

Query Match 99.98; Score 2912; DB 3; Length 575;
Best Local Similarity 99.88; Pred. No. 8.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGVTVLGLALAGLGFPPAPAEPPQSGSCVEHDCFFALYGPATFLNASQICDGLRGHLM 60
DB 1 MGVTVLGLALAGLGFPPAPAEPPQSGSCVEHDCFFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVADVLSLLNGGCGVGRRLWIGLQLPFGCGDPKRLGPRGFQWNTGDNNTSYS 120
DB 61 TVRSSVADVLSLLNGGCGVGRRLWIGLQLPFGCGDPKRLGPRGFQWNTGDNNTSYS 120
QY 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEEOCEVKAADGFLCEFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEEOCEVKAADGFLCEFHFPATCRPLAV 180
QY 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 240
DB 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILPEPSPQRCVNTGGGFCGCHYPNYDLVDG 360
DB 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILPEPSPQRCVNTGGGFCGCHYPNYDLVDG 360
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPANGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPANGLVHSG 516

```

```

RESULT 8
US-09-948-933-228
: Sequence 228, Application US/09948933
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: C1000787
: CURRENT APPLICATION NUMBER: US/09/948,933
: CURRENT FILING DATE: 2001-09-10
: PRIOR APPLICATION NUMBER: 60/231,399
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 6404
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 228
: LENGTH: 575
: TYPE: PRT
: ORGANISM: Human
US-09-948-933-228

```

```

Query Match 99.98; Score 2912; DB 23; Length 575;
Best Local Similarity 99.88; Pred. No. 8.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGVTVLGLALAGLGFPPAPAEPPQSGSCVEHDCFFALYGPATFLNASQICDGLRGHLM 60
DB 1 MGVTVLGLALAGLGFPPAPAEPPQSGSCVEHDCFFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVADVLSLLNGGCGVGRRLWIGLQLPFGCGDPKRLGPRGFQWNTGDNNTSYS 120
DB 61 TVRSSVADVLSLLNGGCGVGRRLWIGLQLPFGCGDPKRLGPRGFQWNTGDNNTSYS 120
QY 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEEOCEVKAADGFLCEFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEEOCEVKAADGFLCEFHFPATCRPLAV 180
QY 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 240
DB 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILPEPSPQRCVNTGGGFCGCHYPNYDLVDG 360
DB 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILPEPSPQRCVNTGGGFCGCHYPNYDLVDG 360
QY 361 ECVPEVDPCEFRANECYQCPPLNQTSTLYCACEGAAPRPHRNRQMCNCTACPADCDPN 420
DB 361 ECVPEVDPCEFRANECYQCPPLNQTSTLYCACEGAAPRPHRNRQMCNCTACPADCDPN 420
QY 421 TOASCCEPEGYIIDDGFTCTIDCEGNGGFCGCHNLPGTFECIGPDSALVRIHTDTC 480
DB 421 TOASCCEPEGYIIDDGFTCTIDCEGNGGFCGCHNLPGTFECIGPDSALVRIHTDTC 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPANGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPANGLVHSG 516

```

Db 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516

RESULT 9

US-09-948-933-318
; Sequence 318, Application US/09948933
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL000787
; CURRENT APPLICATION NUMBER: US/09/948,933
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,399
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 6404
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 318
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Human
US-09-948-933-318

Query Match 99.9%; Score 2912; DB 23; Length 682;
Best Local Similarity 99.8%; Pred. No. 1e-192;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLYIGALALAGLGPAPAEPOGSGQVEHDCFLALPGPATFLNASQICDGLRGLHM 60
Db 108 MGVLYIGALALAGLGPAPAEPOGSGQVEHDCFLALPGPATFLNASQICDGLRGLHM 167
QY 61 TVRSSVADYISILLNDGSGVRRRLMIGQLPGCGDKRLGRLGFWMTGDNNTSYS 120
Db 168 TVRSSVADYISILLNDGSGVRRRLMIGQLPGCGDKRLGRLGFWMTGDNNTSYS 227
QY 121 RMARLDINGAPLCGLCAVSAEAATVPSEPIWEEQCEKADGFLCEHFHPATCRPLAY 180
Db 228 RMARLDINGAPLCGLCAVSAEAATVPSEPIWEEQCEKADGFLCEHFHPATCRPLAY 287
QY 181 EPGAAAASITGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 240
Db 288 EPGAAAASITGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAP 347
QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPP 300
Db 348 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPP 407
QY 301 DQSGSYSCMETGYRLAADHRCEDVDCTLEPSPQRCVNTQGGFEGCHCPNVDLVYG 360
Db 408 DQSGSYSCMETGYRLAADHRCEDVDCTLEPSPQRCVNTQGGFEGCHCPNVDLVYG 467
QY 361 ECEVPVDFRANCEYCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
Db 468 ECEVPVDFRANCEYCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 527
QY 421 TQASCEBPEGYIIDDGFTCTDIDECENGFCGVCNHLPGTEFCICGPDALVRHIGTDC 480
Db 528 TQASCEBPEGYIIDDGFTCTDIDECENGFCGVCNHLPGTEFCICGPDALVRHIGTDC 587
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516
Db 588 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 623

RESULT 10

US-09-331-793-4
; Sequence 4, Application US/09331793
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi

; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/09/331,793
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-331-793-4

Query Match 96.8%; Score 2824; DB 17; Length 497;
Best Local Similarity 100.0%; Pred. No. 8.8e-187;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQSGSCVEHDCFLALPGPATFLNASQICDGLRGLMTRVSSVADYISILLND 78
Db 1 APAEPQSGSCVEHDCFLALPGPATFLNASQICDGLRGLMTRVSSVADYISILLND 60
QY 79 GGVRRRLMIGLOLPFGCGDKRLGRLGFWMTGDNNTSYSRMARLDINGAPLCGLCV 138
Db 61 GGVRRRLMIGLOLPFGCGDKRLGRLGFWMTGDNNTSYSRMARLDINGAPLCGLCV 120
QY 139 AVSAEAATVPSEPIWEEQCEKADGFLCEHFHPATCRPLAYEPGAAAASITGTTPFA 198
Db 121 AVSAEAATVPSEPIWEEQCEKADGFLCEHFHPATCRPLAYEPGAAAASITGTTPFA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAPAMDCSVENGCEHACNA 258
Db 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREAPAMDCSVENGCEHACNA 240
QY 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPPQPSYSCMETGYRLAA 318
Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPPQPSYSCMETGYRLAA 300
QY 319 DQHRCEVDCTLEPSPQRCVNTQGGFEGCHCPNVDLVYGCEVPVDFRANCEYOC 378
Db 301 DQHRCEVDCTLEPSPQRCVNTQGGFEGCHCPNVDLVYGCEVPVDFRANCEYOC 360
QY 379 QPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNQASCEBPEGYIIDDGFT 438
Db 361 QPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNQASCEBPEGYIIDDGFT 420
QY 439 CTDIDECENGFCGVCNHLPGTEFCICGPDALVRHIGTDCSGKVDGDSGSGEPPSP 498
Db 421 CTDIDECENGFCGVCNHLPGTEFCICGPDALVRHIGTDCSGKVDGDSGSGEPPSP 480
QY 499 PTGSTLTPPAVGLVHSG 515
Db 481 PTGSTLTPPAVGLVHSG 497

RESULT 11

US-07-796-336A-14
; Sequence 14, Application US/07796336A

; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: P.C.
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/796.336A
FILING DATE: 19911122
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-066-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-796-336A-14

Query Match 95.0%; Score 2770; DB 3; Length 494;
Best Local Similarity 99.4%; Pred. No. 4.8e-183;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVLVGLALAGLGFPAAPAPPOPGSGQCYEHDCFFALYPPAPFTLNASQICDGLRCHLM 60
DB 1 MGVLVGLALAGLGFPAAPAPPOPGSGQCYEHDCFFALYPPAPFTLNASQICDGLRCHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVGRRLMTGLQLPCCGDPKRLGPRFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVGRRLMTGLQLPCCGDPKRLGPRFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLCGPLCAVSAEAETVSEPIWEBOCEVADGFLCEFHPPATCRPLAV 180
DB 121 RMARLDLNGAPLCGPLCAVSAEAETVSEPIWEBOCEVADGFLCEFHPPATCRPLAV 180
QY 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
DB 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEGTGRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECEVPVDFPCFRANCEYOCOPLNQTSYLCVCAEGFAPIDHEPHRCOMFCNQACPADCPN 420
DB 361 ECEVPVDFPCFRANCEYOCOPLNQTSYLCVCAEGFAPIDHEPHRCOMFCNQACPADCPN 420
QY 421 TQASCECPRGYILDDGFTCTDIDCEENGFCGVCVHNLPGEFECICGDSALVHHTIDC 480
DB 421 TQASCECPRGYILDDGFTCTDIDCEENGFCGVCVHNLPGEFECICGDSALVHHTIDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDGDSGSGE 494

RESULT 12
US-07-796-336A-16
Sequence 16, Application US/07796336A
GENERAL INFORMATION:
APPLICANT: Dol, Takeshi
APPLICANT: Iwasaki, Akio

APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Ohkuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: P. C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/796.336A
FILING DATE: 19911122
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-066-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-796-336A-16

Query Match 94.9%; Score 2768; DB 3; Length 494;
Best Local Similarity 99.2%; Pred. No. 6.5e-183;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVLVGLALAGLGFPAAPAPPOPGSGQCYEHDCFFALYPPAPFTLNASQICDGLRCHLM 60
DB 1 MGVLVGLALAGLGFPAAPAPPOPGSGQCYEHDCFFALYPPAPFTLNASQICDGLRCHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVGRRLMTGLQLPCCGDPKRLGPRFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVGRRLMTGLQLPCCGDPKRLGPRFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLCGPLCAVSAEAETVSEPIWEBOCEVADGFLCEFHPPATCRPLAV 180
DB 121 RMARLDLNGAPLCGPLCAVSAEAETVSEPIWEBOCEVADGFLCEFHPPATCRPLAV 180
QY 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
DB 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEGTGRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECEVPVDFPCFRANCEYOCOPLNQTSYLCVCAEGFAPIDHEPHRCOMFCNQACPADCPN 420
DB 361 ECEVPVDFPCFRANCEYOCOPLNQTSYLCVCAEGFAPIDHEPHRCOMFCNQACPADCPN 420
QY 421 TQASCECPRGYILDDGFTCTDIDCEENGFCGVCVHNLPGEFECICGDSALVHHTIDC 480
DB 421 TQASCECPRGYILDDGFTCTDIDCEENGFCGVCVHNLPGEFECICGDSALVHHTIDC 480

Db 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSGVCCHNLPTRETCICGPDALVRIHGTDC 480
QY 481 DSGKVDGDSGGE 494
Db 481 DSGKVDDEASGSD 494

RESULT 13

US-08-212-384-4
Sequence 4, Application US/08212384
GENERAL INFORMATION:
APPLICANT: KIMURA, SHIGERU
APPLICANT: MIZOGUCHI, TOSHIMI
APPLICANT: OKUCHI, MASAO
APPLICANT: DOI, TAKESHI
APPLICANT: IMASAKI, AKIO
TITLE OF INVENTION: PREPARATION OF A THROMBIN-BINDING
SUBSTANCE
TITLE OF INVENTION: SUBSTANCE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHLOM, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
City: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,384
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,691
FILING DATE: 06-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ohlon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-063-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)412-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-212-384-4

Query Match 94.2%; Score 2746; DB 6; Length 486;
Best Local Similarity 99.8%; Pred. No. 2.1e-181;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGVIVLGLALAGLIGFPAPAEPOPGSGQVEHDFALYPGATFLNASQIDGRLGHLM 60
Db 1 MGVIVLGLALAGLIGFPAPAEPOPGSGQVEHDFALYPGATFLNASQIDGRLGHLM 60
QY 61 TYRSSVADVISLLNDGSGVGRRLMTIGQLPPGCGDKRLGLPGFOWVTGDNNTSYS 120
Db 61 TYRSSVADVISLLNDGSGVGRRLMTIGQLPPGCGDKRLGLPGFOWVTGDNNTSYS 120
QY 121 RWRARDLNGAPLGLCVAAVAEATVSEPIWEEOCEVKADGFLCEHFHPATGRPLAV 180
Db 121 RWRARDLNGAPLGLCVAAVAEATVSEPIWEEOCEVKADGFLCEHFHPATGRPLAV 180
QY 181 EPGAAAAAVSTYGTTPFAARGADFQALPVGSSAAVAAPLGIQLMCTAPGAOVGHHAREAP 240

Db 181 EPGAAAAAVSTYGTTPFAARGADFQALPVGSSAAVAAPLGIQLMCTAPGAOVGHHAREAP 240
QY 241 GAMDCSVENGCEBACNAIPGAPRCQCPAGAAIADGRSCTASATOSCDNDCEHFCVNP 300
Db 241 GAMDCSVENGCEBACNAIPGAPRCQCPAGAAIADGRSCTASATOSCDNDCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLADQHRCEVDVDCILPEPFCORCVNTGGGFECHCYPNYDLVVG 360
Db 301 DPGSYSCMCETGYRLADQHRCEVDVDCILPEPFCORCVNTGGGFECHCYPNYDLVVG 360
QY 361 ECVEPVDPCEFRANCEYOCOPINOTSYLCVCAEGFAP1PHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPCEFRANCEYOCOPINOTSYLCVCAEGFAP1PHEPHRCQMFNCQTACPADCDPN 420
QY 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSGVCCHNLPTRETCICGPDALVRIHGTDC 480
Db 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSGVCCHNLPTRETCICGPDALVRIHGTDC 480
QY 481 DSGKVD 486
Db 481 DSGKVD 486

RESULT 14

US-07-835-436C-1
Sequence 1, Application US/07835436C
GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIEF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/835,436C
FILING DATE: 26-FEB-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OLIEF, JAMES A.
REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 27706
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PTO ALEX
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-835-436C-1

Query Match 92.2%; Score 2690; DB 3; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e-177;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MGVLVGLALAGLGFPAAPAPQSGSCVEHDCFALYPGATFLNASQICDGLRGHLM 60
DB 1 MGVLVGLALAGLGFPAAPAPQSGSCVEHDCFALYPGATFLNASQICDGLRGHLM 60
OY 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPKCGDPRKRLGFGFWTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPKCGDPRKRLGFGFWTGDNNTSYS 120
OY 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEBOQCEVKADGFLCEFHFPATCRPLAY 180
DB 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEBOQCEVKADGFLCEFHFPATCRPLAY 180
OY 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOGHMAAREAP 240
DB 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOGHMAAREAP 240
OY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDCLEHFCVPPN 300
DB 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDCLEHFCVPPN 300
OY 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
OY 361 ECEVPDPCFRANCEYQCCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
OY 421 TQASCEPBGYIIDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDALVRH 475
DB 421 TQASCEPBGYIIDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDALVRH 475

```

```

RESULT 15
US-07-835-436-1
; Sequence 1, Application US/07835436
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVE & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,436
; FILING DATE: 19920327
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIVE, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-07-835-436-1

Query Match          92.1%; Score 2687; DB 3; Length 475;
Best Local Similarity 99.8%; Pred. No. 2.5e-177;
Matches 474; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGVLVGLALAGLGFPAAPAPQSGSCVEHDCFALYPGATFLNASQICDGLRGHLM 60
DB 1 MGVLVGLALAGLGFPAAPAPQSGSCVEHDCFALYPGATFLNASQICDGLRGHLM 60
OY 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPKCGDPRKRLGFGFWTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPKCGDPRKRLGFGFWTGDNNTSYS 120
OY 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEBOQCEVKADGFLCEFHFPATCRPLAY 180
DB 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEBOQCEVKADGFLCEFHFPATCRPLAY 180
OY 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOGHMAAREAP 240
DB 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOGHMAAREAP 240
OY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDCLEHFCVPPN 300
DB 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDCLEHFCVPPN 300
OY 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
OY 361 ECEVPDPCFRANCEYQCCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
OY 421 TQASCEPBGYIIDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDALVRH 475
DB 421 TQASCEPBGYIIDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDALVRH 475

```

Search completed: May 8, 2002, 12:43:30
Job time: 341 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:44:48 ; Search time 48.4 Seconds
(without alignments)
1559.432 Million cell updates/sec

Title: US-09-509-994-2
Sequence: 1 MGVLVIGALALAGLGFAP.....PSPTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_17.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_fodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2658	91.2	468	4 Q9UC32	Q9UC32 homo sapien
2	1844	63.2	577	11 Q35370	P97883 ratus norv
3	1400.5	48.0	461	11 P97883	Q9NPY3 homo sapien
4	587	20.1	652	4 Q9NPY3	Q9HCU0 homo sapien
5	584.5	20.0	757	4 Q9HCU0	000274 homo sapien
6	584	20.0	652	4 000274	089103 mus musculu
7	547	18.8	644	11 Q89103	Q9E611 ratus norv
8	534.5	18.3	643	11 Q9E611	Q9J126 ratus norv
9	532.5	18.3	643	11 Q9J126	Q9TVQ2 caenorhabd
10	431	14.8	1664	5 Q9TVQ2	Q88281 ratus norv
11	414	14.2	1574	11 Q88281	Q9V889 drosophila
12	386	13.2	1394	5 Q9V889	P87363 gallus gall
13	376	12.9	708	13 P87363	Q9WU19 ratus norv
14	375	12.9	2906	11 Q9WU19	Q88840 mus musculu
15	367	12.6	3857	11 Q88840	Q9WU18 ratus norv
16	361	12.4	2872	11 Q9WU18	Q9CX88 mus musculu
17	359	12.3	528	11 Q9CX88	Q99K58 mus musculu
18	358.5	12.3	1174	11 Q99K58	Q9NS21 homo sapien
19	357	12.2	741	4 Q9NS21	

20	353.5	12.1	1242	4 Q9NS15	Q9NS15 homo sapien
21	353.5	12.1	1382	4 Q9H7K2	Q9H7K2 homo sapien
22	351.5	12.1	1511	4 Q75412	Q75412 homo sapien
23	351.5	12.1	1587	4 Q00508	Q00508 homo sapien
24	341	11.7	576	4 Q9Y3V7	Q9Y3V7 homo sapien
25	340	11.7	1062	11 Q60789	Q60789 mus musculu
26	333.5	11.4	961	11 Q9E0C6	Q9E0C6 mus musculu
27	332.5	11.4	1764	11 Q35806	Q35806 ratus norv
28	332.5	11.4	1964	11 Q35442	Q35442 mus musculu
29	332	11.4	999	4 Q9NQ36	Q9NQ36 homo sapien
30	331.5	11.4	1713	11 Q88349	Q88349 mus musculu
31	331	11.4	589	5 Q9T2S1	Q9T2S1 caenorhabd
32	329	11.3	2189	5 Q9B105	Q9B105 elmeria ten
33	328.5	11.3	956	11 Q99K64	Q99K64 mus musculu
34	325	11.1	1095	11 Q60784	Q60784 mus musculu
35	324.5	11.1	937	5 Q9BLJ1	Q9BLJ1 ciona intes
36	323	11.1	1253	11 Q61810	Q61810 mus musculu
37	321	11.0	2321	4 Q9Y6L8	Q9Y6L8 homo sapien
38	321	11.0	2321	4 Q9UM47	Q9UM47 homo sapien
39	319.5	11.0	1833	11 Q08999	Q08999 mus musculu
40	316	10.8	2281	4 Q9PPL3	Q9PPL3 homo sapien
41	314.5	10.8	3507	5 Q23587	Q23587 caenorhabd
42	314	10.8	997	11 Q9JUS0	Q9JUS0 mus musculu
43	313	10.7	798	5 Q18026	Q18026 caenorhabd
44	312.5	10.7	1821	4 Q14767	Q14767 homo sapien
45	312	10.7	517	4 Q9NP01	Q9NP01 homo sapien

ALIGNMENTS

RESULT 1
ID Q9UC32 PRELIMINARY; PRT; 468 AA.
AC Q9UC32;
DT 01-MAY-2000 (TRMBLrel. 13, Created)
DT 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93293792; PubMed=8390446;
RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;
RT "Primary thrombomodulin, its isolation and characterization.";
RL J. Biochem. 113:433-440(1993).
DR HSSP: P07204; 12AQ.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001304; lectin-C.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF; 5.
DR Pfam: PF00059; lectin-C; 1.
DR PRINTS: PR00907; THROMBOMODUL.
DR SMART: SM0034; CLECT; 1.
DR SMART: SM00181; EGF; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS50041; C-TYPE-LECTIN-2; 1.
DR PROSITE: PS01186; EGF-2; 2.
DR PROSITE: PS01187; EGF-Ca; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 468 AA; 49444 MW; 4BFE8E9BFB86A40 CRC64;

Query Match 91.2%; Score 2658; DB 4; Length 468;
Best Local Similarity 99.8%; Pred. No. 4.4e-204;
Matches 467; Conservative 1; Indels 0; Gaps 0;
QY 19 AFAEPGSGQVEHDCFAFLYGPATFLNASQICDGLRGLMTVRSVADYISLLAND 78

```

Db 1 APAEPQPGSGOCVEHDCALYPGPATFLMASQICDGLGHMLTVRSSVAADVISILLNGD 60
QY 79 GGVRRRLMIGIQLPPGGDKRLGLPLRGFWMTGDNNTSYSRMARLDLNGAPLCPLCY 138
Db 61 GGVRRRLMIGIQLPPGGDKRLGLPLRGFWMTGDNNTSYSRMARLDLNGAPLCPLCY 120
QY 139 AVSAEAATVPSEPIWEEOCEVKADGFLCEFHFPATCRPLAVEGAAAAYSTYGPFA 198
Db 121 AVSAEAATVPSEPIWEEOCEVKADGFLCEFHFPATCRPLAVEGAAAAYSTYGPFA 180
QY 199 ARGADFOALPVSSSAVAAPLGLQMLCTAPPAVAGVGHMAREAPGAMDCSVENGGEHCNA 258
Db 181 ARGADFOALPVSSSAVAAPLGLQMLCTAPPAVAGVGHMAREAPGAMDCSVENGGEHCNA 240
QY 259 IPGARPCCPGGAALQAGRSCTASATQSCNDLCEHFCVPMPDPGYSCECTGYRLAA 318
Db 241 IPGARPCCPGGAALQAGRSCTASATQSCNDLCEHFCVPMPDPGYSCECTGYRLAA 300
QY 319 DOHRCEDVDCTLEPSPCPCRCVNTGGGFECCHCPNNDLVDECEVEPYDPCFRANCEYOC 378
Db 301 DOHRCEDVDCTLEPSPCPCRCVNTGGGFECCHCPNNDLVDECEVEPYDPCFRANCEYOC 360
QY 379 OPLNOTSYLCVCAEGFAPIRPHPRHRCMFCNOTACPADCDNTQASCSCEPGYTLDDGFI 438
Db 361 OPLNOTSYLCVCAEGFAPIRPHPRHRCMFCNOTACPADCDNTQASCSCEPGYTLDDGFI 420
QY 439 CTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHIGTDCDSGKYD 486
Db 421 CTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHIGTDCDSGKYD 468

RESULT 2
035370 PRELIMINARY; PRT; 577 AA.
AC 035370:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RN Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1; -
DR EMBL; AF022742; AAB80923.1; -
DR HSSP; P07204; IFGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; lectin_C.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF_5.
DR Pfam; PF00059; lectin_C_1.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA_2.
DR SMART; SM00001; EGF_Like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

```

```

SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;
Query Match 63.2%; Score 1844; DB 11; length 577;
Best Local Similarity 64.0%; Pred. No. 4, 1e-139;
Matches 333; Conservative 49; Mismatches 132; Indels 6; Gaps 5;

QY 1 MGVVLVGLALATLGFPAEPQPGSGOCVEHDCALYPGPATFLMASQICDGLGHML 60
Db 1 MGVVLVGLALATLGFPAEPQPGSGOCVEHDCALYPGPATFLMASQICDGLGHML 60
QY 61 TVSSVAADVISILLNGDGGVRRRLMIGIQLPPGGDKRLGLPLRGFWMTGDNNTSY 120
Db 61 TVSSVAADVISILLNGDGGVRRRLMIGIQLPPGGDKRLGLPLRGFWMTGDNNTSY 118
QY 121 RMARLDLNGAPLCPGLCVAVSAEAATVPSEPIWEEOCEVKADGFLCEFHFPATCRPLAV 180
Db 119 RMARLDLNGAPLCPGLCVAVSAEAATVPSEPIWEEOCEVKADGFLCEFHFPATCRPLAV 178
QY 181 EP-GAAAAYSTYGPFAARGADFOALPVSSSAVAAPLGLQMLCTAPPAVAGVGHMARE 239
Db 179 NTRDPEGAHISSTYNTPLVGSADFOTLPIGSSATVAPGLEIVCRALPGTSEGHWTREV 238
QY 240 PGAMDCSVENGGEHCNAIPGARPCCPGGAALQAGRSCTASATQSCNDLCEHFCVPN 299
Db 239 TGAMDCSVENGGEHCNRSANGRCVCSGGLQADGHSCKAPVAQLCNELQHCNVNN 298
QY 300 PDGPGYSCECTGYRLAADOHRCEDVDCTLEPSPCPCRCVNTGGGFECCHCPNNDLV 359
Db 299 SDVPGYSCECTGYRLAADOHRCEDVDCTLEPSPCPCRCVNTGGGFECCHCPNNDLV 358
QY 360 GECVEPYDPCFRANCEYOCPLNOTSYLCVCAEGFAPIRPHPRHRCMFCNOTACPADCDP 419
Db 359 GECVEPYDPCFRANCEYOCPLNOTSYLCVCAEGFAPIRPHPRHRCMFCNOTACPADCDP 418
QY 420 NTQASCSCEPGYTLDDGFICTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHIGT 479
Db 419 NSPFCQPCFGLDGSCTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHIGT 478
QY 480 CDSGKY--DGDGSGSGEPSPS-PTPGSTLPPPAVGLVHSG 516
Db 479 CDPVLEDESDGSGEPSPSNTPTVSSIVPPPARMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
AC P97883:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE THROMBOMODULIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RN Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90121; AAB49723.1; -
DR HSSP; P07204; IFGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF_5.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_Like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 3.

```


RL J. Biol. Chem. 276:7408-7414(2001).
 DR EMBL: AF279142; AAG00867.1; -
 DR EMBL: AJ295846; CAC34381.1; -
 DR InterPro: IPR001152; Asx_hydroxyl.
 DR InterPro: IPR005561; EGF-like.
 DR InterPro: IPR01881; EGF-Ca.
 DR InterPro: IPR01304; lectin_c.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00059; lectin_c.1.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00181; EGF_3.
 DR SMART: SM00179; EGF_Ca_2.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
 DR Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 757 TUMOR ENDOTHELIAL MARKER 1.
 SQ SEQUENCE 757 AA; 80859 MW; C96363EAI1FD8FFA0 CRC64;

Query Match 20.0%; Score 584.5; DB 4; Length 757;
 Best Local Similarity 33.6%; Pred. No. 1,3e-38;
 Matches 177; Conservative 44; Mismatches 191; Indels 115; Gaps 24;

QY 1 MGVLVGLAALALAGLGFAPAPAPQPGSGQVCHDPCFALPGPATFLNASQICDGLRGHL 59
 DB 2 LKLLILMAAAGPTLQODPMAAEPR--AACGPGSCYALFPRRTFLEAWRACRELGDGL 58
 QY 60 MYRVSSVAADVIVSLINGDGVGRRRLMIGLQLPFGGDPKRLGRLPGFWMTGDNNTSY 119
 DB 59 APTPTREEAQRVDLVG--AGPASKRLMIGLQARQCOLQR--PLKGFMTGDDDTAR 114
 QY 120 SRMARLDLNGARLCPGLCYAVSAEAATVSEPTWEEQCEVKADGFLCEHPFATCPPLA 179
 DB 115 TTNAAQ-PASGGPCPAPQRCVLALEAS--GEHRLGSGCLTAVDYLQCGFEGGACPALQ 168
 QY 180 VERGAAAAVSTYTGPPFAPARGADPQALRYGSSAAV--APGLQIMCTAPRGAVQGHMA 236
 DB 169 DEAGQAGPAPV---YTPPFILVSTEFEMLPFGSYAAVQCAGRGASLICYAPREGVYG-MS 224
 QY 237 REAP---GAMDCSVENGGEHAC-NAIPGAPRCQAGALADGSSCTASATQS-CNDL 231
 DB 225 RAGPLCLGT-GSPDNGGCEHECEVEVDHVSCTEGSRFLADGSSCEDPCAACDEQ 283
 QY 292 CEHFVCPNPDQPGSYSCMCEGTGRLAADQ-HRCEDVDCCILEPSPQRCVNTQGGFECH 350
 DB 284 CE-----PGFPQYSGCHCRLGFRPAEDDPHRCVDTDEQI-AGVQCOQCVNTVGGFECH 336
 QY 351 CYPNVDLVGCEVEPVPDPCFRANCEYQCPPLNQTSTLYCVAEGFADIPHEPRQCMFCNQ 410
 DB 337 CSEGHLE-----ADGIS----- 349
 QY 411 TACPADCDP---NQASCEPEGYIIDDGFICTDIDE--GENGFGCGVCHNLPGT-- 461
 DB 350 -----CSPAGMAGASQDLDG-ELDDGDEDEDEDMKAFNGMT-----EMPGILM 397
 QY 462 FEGICGPDALARIHIGTDCDSKVDGSGSEPPSPPGSTLTPP 508
 DB 398 MEPTPPDFALAYRPSFPD-----REFQIPIPEPTWPP 432
 RESULT 6
 000274 PRELIMINARY; PRT; 652 AA.
 AC 000274;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE C10R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 QX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97199258; PubMed=9047234;
 RA Neomunceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
 RT "cDNA cloning and primary structure analysis of C1qR(P), the human
 C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
 RL Immunity 6:119-129(1997).
 DR EMBL: U94333; AAB53110.1; -
 DR HSSP; P35555; 1EMN.
 DR InterPro: IPR001152; Asx_hydroxyl.
 DR InterPro: IPR005561; EGF-like.
 DR InterPro: IPR01881; EGF_Ca.
 DR InterPro: IPR01304; lectin_c.
 DR Pfam: PF00008; EGF_5.
 DR Pfam: PF00059; lectin_c.1.
 DR SMART: SM00034; CLECT_1.
 DR SMART: SM00179; EGF_Ca_3.
 DR SMART: SM00001; EGF_Like_2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_Ca_3.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 652 AA; 68576 MW; B7EAB5FE5714A775 CRC64;

Query Match 20.0%; Score 584; DB 4; Length 652;
 Best Local Similarity 31.8%; Pred. No. 1,2e-38;
 Matches 181; Conservative 56; Mismatches 209; Indels 124; Gaps 32;

QY 2 LGVLVGLAALALAGLGFAPAPAPQPGSGQVCHDPCFALPGPATFLNASQICDGLRGHL 59
 DB 5 MGILLILLILLIPGAGTGADTE---AVYCVGTACTYHAGSLSAEAQNHQNGNL 60
 QY 60 MYRVSSVAAD---VISLLINGDGVGRR--RLMIGLQLPFG--CGDPKRLGRLPGFWMT 112
 DB 61 AIVKSKKEAHOYORVLAQLRLREALTARMSKFMIGLQREKGCIDPSL--PLKGSWVG 118
 QY 113 GDNNTYSRRARLDLNGARLCPGLCYA--VSAEAATVPSB-PIWEEQOC-----EVKAD 163
 DB 119 GGEDETPYSNMHKLRLNSC--ISKRCVSLDLISQPLLPNLPMWSBSPGSGSPSPSNI 176
 QY 164 GFLCEHPFATCPPLAV-EPGAAAAVSTYTGPPFAPARGADPQALRYGSSAAVAPGLQL 222
 DB 177 GFYCKSFYKCMCPRLALGPG-----QVYITTFQTTSSLEAVFPASANVA----- 224
 QY 223 MC-TAPRGAVQGH---AREAPGAMD-----CSYENGCEHAC-NAIPGAP 263
 DB 225 -CGEADKDETHSYFLICEKAPDVPFDMGSSGPLCVSPKGCNFMNNGCHODCEGGDGSF 283
 QY 264 RCQCPAGALADGRCSTASATQSCNDLCE--HFCVPNPDQPGSYSCMCEGTGRLAADQH 321
 DB 284 LCGCRPGFRLDLDTVTC-ASRNPCSSPCRGATCVLGP-HGKNYTCRCPOGYLDSOL 341
 QY 322 RCEDVDCCILEPSPQRCVNTQGGFECHCYRYDLVDECEVPVPDPCFRANCEYQCP 381
 DB 342 DCYVDVDEC--QDSPCAQCECVNTPGFRCEQWGY-----EPGSP-----GEGACQDV 386
 QY 382 NQTSYLCVCAEGFAPIPHEPRHRCQMFQNOTACPADCDPNTQAS--CECEGYIL--DDGF 437
 DB 387 DE-----CALGRSP-----CAQGC-TTIDGSPFCSCGEGYVLAGEDET 423
 QY 438 ICTDIDEC--ENGFGCGVCHNLPGTFCICGPDALARIHIGTDCDSKGV-----D 486
 DB 424 QCDVDCEVCGPGLCDSCFNTQGSFHCGLPGWTLAN-GVSCMTMGVSLGPPSGPFD 482
 QY 487 GGDSSGE-----PPSPPTPGSTLTPPA 509
 DB 483 EEDKGEKESIVPRAATASPTRGPESTPKA 512

RESULT 7
 089103

Query Match	18.8%	Score 547	DB 11	Length 644
Best Local Similarity	30.2%	Pred. No. 1,1e-35		
Matches	168	Conservative	55	Mismatches 210
				Indels 124
				Gaps 27

QY	5	LVGLHALALAGIGPPAPAEPOGGSGVEHDFCALYRGPATFLMNSQICDGLRGLHMYVNS	64
		:	:
DB	9	LLGLLGLQPMWG--AAADSG--AVVECGTACTAHMGKLSIAAEQHRKENNGSLATVKS	64
		:	:
QY	65	SVAA-----DVSLTLINDGGVGRR--RLWIGIQLEPPGGDPRKRLPLRGEFOWYTGDNNTS	118
		:	:
DB	65	EEEAHWQOALITQLTKFKAPLEAKMKGFWTIGFOREKNGCTYHDL--PMRGFSWVGGGEDTA	123
		:	:
QY	119	YSRRARLDINAPALCGSLCAVAASAATATPSE--PIREBOGCV-----KADGFLCEHF	171
		:	:
DB	124	YSNMYKASKSCICIRKCVSLITDLSLCTPHPSHLPKHESPGCTPPAPONSITGFLCKENF	183
		:	:
QY	172	PATCRPLAV--EPGAAAAAVSITVGTTPPAARGADFDALPVGSSAAVAPLGIO-----LM	223
		:	:

Dn	184	KNCGRPLALGPR-----RVTYTPPQATTSLSLEAVPPASVANVA--CGDEAKSETHYL	236
Oy	224	CT-APPGAVOGHWRARAP----GAMDCSVENGCEHNAC-NATIGCARPCOCFAGALQADG	277
Dn	237	CNEKTRPIF--HWSSSGPLCVSPKKGCSFNNGGCQQDCEFGDDSFRCGRPEFRLLDL	294
Oy	278	RSCSTASTOSCNLDLCHFCVCPNDDPG-----SYSOMCETGYRLAADGHCERYD	327
Dn	295	VTCAS-----RNPCSSNRPCRGGGMCHSVPLSENTRYCRCPGSYGQLDSSGVHVDID	344
Oy	328	DCILIEPSPCFORCVNTQGFECHCIPTNYDLVG--ECVEPYDPCFANREYCCPLNPT	384
Dn	345	EC--QDSPCADQDVNLTSFHCWCWVYO-PSGKEENCEVDCAANAHP-----	392
Oy	385	SYLVCACEGFAPIPHEPHRCOMFCNOTACPADCDPNTQAS--CECPGYTL--DDGEICT	440
Dn	393	-----CAQGGI-----NNDGPSYCCKRGYIVSSEDSTQOE	423
Oy	441	DIBECEN--GGFSGCYHNLPGTFEFCICGPBSALARHGIDCDSG-----KVD	486
Dn	424	DIEDCSARANKPDSCJCFNTDSSFRCGCGPWELAPN-GVFCSHGTIVSELPARPKOED	482
Oy	487	GSDSGSGEPSPPTPGS	503
Dn	483	NDRKESTMPTMPSS	499
RESULT	8		
ID	09ET61	PRELIMINARY:	PRF: 643 AA.
AC	09ET61		
DT	01-MAR-2001 (Tremblrel, 16, Created)		
DT	01-MAR-2001 (Tremblrel, 16, Last sequence update)		
DT	01-JUN-2001 (Tremblrel, 17, last annotation update)		
DE	CIO/MBL/SPA RECEPTOR CLORP PRECURSOR.		
GN	CLORP.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=101116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-PVG:		
RA	Iovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.,		
RT	"Characterization and molecular cloning of rat ClqBP, a receptor on		
RL	macrophages, natural killer cells and dendritic cells."		
EMBL	Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.		
DR	EMBL: AF136537; AAg01572.1; "		
DR	InterPro: IPR000152; Asx_hydrolyl.		
DR	InterPro: IPR000561; EGF-like.		
DR	InterPro: IPR001881; EGF Ca.		
DR	InterPro: IPR001304; lectin_c.		
DR	Pfam: PR00008; EGF_5		
DR	SMART: SM00034; CLECT_1.		
DR	SMART: SM00181; EGF_5.		
DR	SMART: SM00179; EGF_CA_5.		
DR	SMART: SM00001; EGF_like_2.		
DR	PROSITE: PS00010; ASX_HYDROLYL; UNKNOWN_1.		
DR	PROSITE: PSS0041; C_TYPE_LECTIN_2; 1.		
KW	Signal; Receptor.		
FT	SIGNAL 1	POTENTIAL.	
SEQ	SEQUENCE 643 AA: 68781 MW: 9A4C933AD943DB6 CRC64:		
Query Match	18.3%; Score 534.5; DB 11; Length 643;		
Best Local Similarity	29.9%; Pred. No. 1.1e-34;		
Matches 167; Conservative 56; Mismatches 208; Indels 127; Gaps			
Oy	4 VLVIGALALAGLCPPPAEAPQPGSGSCVEHDCEFALYPGPATFIYNAISOICGLRGHTMTVR	63	
Dn	8 LILLGLLGQLMAGAANDSE-----AYVCEGTACTITTHMKGLSAERQRHCRNENSGNLAIVK	63	
Oy	64 SSVA-----DIVISILLN---GDGGVGRRRRIWGQLDPPCGDPRKLGPLRGFWWTGDN	115	

```

Db 64 SEERHVOEALADLLKTKKASEETKIG--KFWIGLOREKCKTYHDL--PMKGFVWVGGE 120
QY 116 NTSYRMARLDLNG--APLCGPLCYAASAAEATVPSE--PIWEPOCEV-----KADGRL 166
Db 121 DTTYSNMWKASKSCISKRCVSLIDLSLKRH--PSHLPRMHSFGCTDPAAGNSTIEGRL 178
QY 167 CEFHPATCRPLAV--EPGAAAANVSTYGPFAAGADFOALPYGSSAAVAPLIGLQ---- 221
Db 179 CKFNFKGMSPLALGPG-----QLTYTTPFOATTSLSKAVFAVAVV--CGDEAESK 231
QY 222 ---LMCTAPPAVQGHMAREAP---GAMDCSVNGGCEHAC--NAIPGAPROCCPAGAL 273
Db 232 TNYLCKRTTAGV--FHWSSGSLCYSPKFGCSFNNGGCCODCEGGDGSFRGCGRGRRL 290
QY 274 QADGSCSTASATQSCNDLCEHFVCPNDPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTGAS-----RNPCCSNPCTGGGCMCHSVPLSENCTHCPRGYOLDSSQVHC 340
QY 324 EDVDDCILEPSPORCVNTGSGFCHCYPNYDLVDG--ECVFPVDPCCFRANCEYQOQPL 381
Db 341 VDIDEC--EDSPCDOECINTPGFHCFCWGYOSSGSKERACEVDDEC----- 386
QY 362 NOTSYLCVCAEGFAPIPHEPRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGF 437
Db 387 --TAAYSPCAOG-----CT-----NTDGSFYSCKEGYIMSGEDST 420
QY 438 ICTDIDECENGSGSVCHNLPGTFECICGPDALRHIGTDCDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRGCGPAGFELAPN--GVSCTRGSMFSELPARPQKE 478
QY 486 DGGDSGSGEPPSPPPGS 503
Db 479 DKDGKESTVPLTEMPGS 496

RESULT 9
Q9J126 PRELIMINARY; PRT: 643 AA.
ID 09J126;
AC 09J126;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE C1ORP.
GN C1ORP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN=MISTER; TISSUE=LUNG;
RA Dean Y.D., McCreath E.P., Akatsu H., Morgan B.P., Gasque P.;
RT "Cloning of the Rat C1ORP cDNA";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF160978; AAF80402.1; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF_Ca.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001304; lectin_C.
DR Pfam: PF00008; EGF_5.
DR SMART: SM00034; C1ECT; 1.
DR SMART: SM00179; EGF_Ca; 3.
DR SMART: SM00001; EGF_Ca; 3.
DR PROSITE: PS00010; Asx_HYDROXYL; 2.
DR PROSITE: PS00041; C_Type_Lectin_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_Ca; 3.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 643 AA; 68780 MW; FB23D742E02A4931 CRC64;

```

Query Match 18.3%; Score 532.5; DB 11; Length 643;

```

Best Local Similarity 29.9%; Pred. No. 1,5e-34;
Matches 167; Conservative 55; Mismatches 209; Indels 127; Gaps 28;

QY 4 VIVIGLALAGLGFAPAPAPQPGSQCEHDCFFALYGPATFTLNAISQIDGLGHMTYR 63
Db 8 LILIGLGLMAGAABDE---AVCEGTACTAHMGKLSAABDORHRCNENGSLATYK 63
QY 64 SSVA--DVISLLN---GDGSGRRRLWIGLQIPCCGDPKRLGPLRGFWYTGDN 115
Db 64 SEERHVOEALADLLKTKKASEETKIG--KFWIGLOREKCKTYHDL--PMKGFVWVGGE 120
QY 116 NTSYRMARLDLNG--APLCGPLCYAASAAEATVPSE--PIWEPOCEV-----KADGRL 166
Db 121 DTTYSNMWKASKSCISKRCVSLIDLSLKRH--PSHLPRMHSFGCTDPAAGNSTIEGRL 178
QY 167 CEFHPATCRPLAV--EPGAAAANVSTYGPFAAGADFOALPYGSSAAVAPLIGLQ---- 221
Db 179 CKFNFKGMSPLALGPG-----QLTYTTPFOATTSLSKAVFAVAVV--CGDEAESK 231
QY 222 ---LMCTAPPAVQGHMAREAP---GAMDCSVNGGCEHAC--NAIPGAPROCCPAGAL 273
Db 232 TNYLCKRTTAGV--FHWSSGSLCYSPKFGCSFNNGGCCODCEGGDGSFRGCGRGRRL 290
QY 274 QADGSCSTASATQSCNDLCEHFVCPNDPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTGAS-----RNPCCSNPCTGGGCMCHSVPLSENCTHCPRGYOLDSSQVHC 340
QY 324 EDVDDCILEPSPORCVNTGSGFCHCYPNYDLVDG--ECVFPVDPCCFRANCEYQOQPL 381
Db 341 VDIDEC--EDSPCDOECINTPGFHCFCWGYOSSGSKERACEVDDEC----- 386
QY 382 NOTSYLCVCAEGFAPIPHEPRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGF 437
Db 387 --TAAYSPCAOG-----CT-----NTDGSFYSCKEGYIMSGEDST 420
QY 438 ICTDIDECENGSGSVCHNLPGTFECICGPDALRHIGTDCDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRGCGPAGFELAPN--GVSCTRGSMFSELPARPQKE 478
QY 486 DGGDSGSGEPPSPPPGS 503
Db 479 DKDGKESTVPLTEMPGS 496

RESULT 10
Q9TYO2 PRELIMINARY; PRT: 1664 AA.
ID Q9TYO2;
AC Q9TYO2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE Y64G10A.7 PROTEIN.
GN Y64G10A.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11;
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
DR SEQUENCE FROM N.A.
DR MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
RN 13;
RP SEQUENCE FROM N.A.
RA Ainscough R.;
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AL117206; CAB60454.1; -

```

QY	119	YSNRARDLNGAPLCPGLVAVSAAE-----ATVSEPIMEPOOCE--VKADGFLCEP	169
Db	56	YLRFAFRSRRGCSKCCILIRQANCSADLCHNGGTCVSEHNDNEQVCECFVGTGAKQY	115
QY	170	HEPATCRPLAVEPGAAAAAASITYGTPEPAARGADFOALPVGSSAAVAPLGLQIMCTAPG	229
Db	116	D-ANEC--MANNGCHEHCVA-TIGTY-----CRCMPG	145
QY	230	AVGHMNAAREAPGANDCSVEENGCGCHCAONATPGAPROCCPAGALQADGRCSTA-----	282
Db	146	FELSGDGTSCSDIDECAASNGGCSDRCVNSGGERCCPDDLYIHAGRGTCGSGFHENL	205
QY	283	---SATOSC---NDLCEHFVCPNPDQPGSYSCMETGYRLAADQHRCEDYDDGCILEPSPC	336
Db	206	ILIKRVTSCTSDNGGCEHEC-ENDSNBEFYRCRCRVGFKLSENRSCQVPDCPDNKGCC	264
QY	337	PQRCVNQGGEECHCYRHYD-VGCEVEPEYDPCEFRAN-CEYQCPQLNQSYLVCYAEGF	354
Db	265	QHCTCNHNGRAOCCQCPGFHLSTDRSCVDDIBECANNKGCHFCENKGT-YRCKCREGY	323
QY	395	APIDHEPRRCOMF-----CNOTACPADC--DPNTQASCCEPEGYIL-DDEFTCTDIDEC-	445
Db	324	-QLGHDGPTCEEMLAGGCQVGGGQHCQCDPDQDGHVCKCKGNGYILANDQKLCHDINECH	382
QY	446	ENGGRCSVCNMLTGTEPCTCIGPDSALARIHGT-----DDDS-----GKVDGDSG	451
Db	383	ENNGDCSDICVNLAGSVCCQCKPGRFLMKDRKTECIDEISSNNGGCEQICSNQSGGYMC	442
QY	492	SGEP	495
Db	443	SCEP	446
RESULT	11		
088281			
ID	088281	PRELIMINARY;	PRT; 1574 AA.
NC	088281		
DT	01-NOV-1998	(TREMBLrel. 08, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	MEGF6.		
GN	MEGF6.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCBI_TaxID=10116;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;		

Query Match	14.28;	Score 414;	DB 11;	Length 1574;
Best Local Similarity	34.18;	Pred. No. 1e-24;		
Matches 104;	Conservative 28;	Mismatches 101;	Indels 72;	Gaps 17;

OY	224	TAPBP--AVOQHMAREARPMADCSVENGSGENHNCMIIPGAPRQSCAGALQADGGSSCA	282
Db	147	CRCPPGQLQDGL-GKTCQDYDECBANHGSCCHRCVNTPGSLDCCKPRFRLNHDGRKTL-	204
OY	283	SATOSC---NDLSEHFVP-----	309
Db	205	-AISSTLGNGGCCHOCVOLVTVOHCRCAPROYOLOEDRGRCVARSPCAGNCGMHIQ	263
OY	310	-----CEGYRLAADHRCEVDYDDCTLEPSPQRCPVNTQGGFECCYPNYDL-VY	359
Db	264	ELRELACGCHPGYQLAADRTKCEDYDVBCALGLACGHLGNLTQGSFKCYCHANGYTELGD	323
OY	360	G-EC-----VEPYDPCFRAN--CEYOCORPLNOTSTLYCAGAGPARIRIPIHRQMC----	408
Db	324	GRQCYRIEMELVNSCGNGSGSHGSH-TSTGLTCTCPRGY-----ELDEDKDTIDID	377
OY	409	---NOTACRPADCDNTQA--SCEEPESYIID-DGFITDIIDECENG--GFCSGVCNNLPPT	461
Db	378	DCANSPPCQQMC-AHTPGYECCSFAGYRLNTDQCGCEDVEDECASGHGCCENHNCNLAS	436
OY	462	FECIC 466	
Db	437	FQCFQ 441	

	RESULT	12
09V589		
ID	09V589	PRELIMINARY; PRT; 1394 AA.
AC	09V589;	
DT	01-MAY-2000 (TREMBLrel. 13)	Created
DT	01-MAY-2000 (TREMBLrel. 13)	Last sequence update
DT	01-JUN-2001 (TREMBLrel. 17)	Last annotation update
DE	CG7526 PROTEIN (FRAGMENT).	
GN	CG7526.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_taxid=7227;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RP	STRAIN-BERKELEY;	
RC	MEDLINE=20196006; PubMed=10731132;	
RX	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Manalidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe W., Pfeiffer B.D.,	

Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Abayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borova D., Botchan M.R., Bouck J., Brokstein P., Brothman P.,
 Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferriz C., Ferreira S., Fleischman W.,
 Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
 Mervulov G., Mishina N.V., Mobarry C., Morris J., Moshnell A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 The genome sequence of Drosophila melanogaster.;
 Science 287:2185-2195(2000).
 RL
 EMBL: AE003538; AAF50536.1; -;
 DR HSSP: P07204; IADX.
 DR
 DR Flybase: FBgn0035798; CG7526.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_12.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00332; CCP_2.
 DR SMART: SM00179; EGF_Ca; 9.
 DR SMART: SM00001; EGF_Like; 5.
 DR PROSITE: PS00010; ASX_HYDROXYL; 8.
 DR PROSITE: PS01186; EGF_2; 9.
 DR PROSITE: PS01187; EGF_Ca; 10.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 FT NON_TER 1394 1394
 SQ SEQUENCE 1394 AA; 152269 MW; CD23380E3162F68A CRC64;

Query Match 13.2%; Score 386; DB 5; Length 1394;
 Best Local Similarity 32.1%; Pred. No. 1.6e-22;
 Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;

224 CTAPGAVG-----HNAKAPGAW-----DCSVNGGCEHNCMAIPG 261
 DB 563 CLCPFGALGNDHNTYLSNFFITDSSTSPSHHTGLDIDECISLNGKSHCQNEPG 622
 QY 262 APRGCPAGALQADGRSCTASATQSC--NDLCENFCVPPNDPGSYSCMCTGYRLAA 318
 DB 623 GFGCACPGLGVALSEDMKTC--QDIDECIDLSNGCSQSLCL--NDGGFACACETGFEITP 677
 QY 319 DQHRCDVDVDCILFSPRCQRCVNTQGGFECGYNYDLDGE--CVPPVPC---FRAN 373
 DB 678 DFGCADIDECSDYXNCSDICINLGHACACERGYELADKLKSLD--VECCGALLSGG 736
 QY 374 CEFQCCPLNQT--SYLCVCEGFAPIRPHPRCOMFCNQTACPA--DCDPPNQASC----- 425

DB 737 CSHEC--INKAGTFEGCGPLGY--ILNDGR-----SCSPALVCGPPGTORSADGAP 785
 QY 426 -ECPEGYILDDGFICTDIDEC--ENGFCGCVCHNLPGTFECICGP-----DSALARI 476
 DB 786 IECNNGYILGSDKCVIDIEQKONGS-CSHRCSNTEGSKCSPGYILDSQTKCQDI 844
 QY 477 GTDCDSK 484
 DB 845 -DECDQK 851

RESULT 13
 ID P87363 PRELIMINARY; PRT; 708 AA.
 AC P87363;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE FIBRILIN-1 (FRAGMENT).
 GN FBN1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Zhou G., Price C., Godfrey M.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U88872; AAB48531.1; -;
 DR HSSP: P07204; IREG.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002212; TB.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_14.
 DR Pfam: PF00683; TB; 2.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00179; EGF_Ca; 14.
 DR PROSITE: PS00010; ASX_HYDROXYL; 13.
 DR PROSITE: PS01186; EGF_2; 10.
 DR PROSITE: PS01187; EGF_Ca; 13.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 FT NON_TER 708 708
 SQ SEQUENCE 708 AA; 76163 MW; C247271C1DF73361 CRC64;

Query Match 12.9%; Score 376; DB 13; Length 708;
 Best Local Similarity 31.8%; Pred. No. 5.1e-22;
 Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

244 DCSVNGGCEHNCMAIPGAPRCOPAGALQADGRSCTASATQSCNDLCENFCVPPNP--- 300
 DB 223 EGSIMNGCENCTGSESTYSCSKQGFALMPHRTCT-----DIDE--CEDNPIC 272
 QY 301 -----DQGSYSCMCTGYRLADQHRCEVDVDCILFSPC--PQRCVNTQGGFECGYR 353
 DB 273 DGGCTNIPGEYRCICLYGFMASEDMDKTCVNVNCDLHPNICISGICENTKGSFICHCMD 332
 QY 354 NYDLDGE--CVPPVPC--FRANCYQCPLN--QTSYLCVCEGRA-----PTP 398
 DB 333 GTSKKKGTTCGD--INBEIGAHMCDRAHVCNTNIPGFSKCSGSGMIGNICKTDLDECS 391
 QY 399 HEPHRCOMFCNQTACPADCDPNTQAS--CECPESYIILDDGFICTDIDE-----CENG- 448
 DB 392 NCTHKSPH-----AC-KNTMGSYRCICKBY--TGDEFCTDIDECSENLNLENGQ 442
 QY 449 -----GF-----CS-----GVCNNDPETFECICGPPSALA 473
 DB 443 CUNAPGYRCEDMGFLPSLDKACEDIDECSLPNICVYGTCHNLDPLFCECEVAYELD 502
 QY 474 RHIG-----TDCDSK--VDGDSGSGEPPSPPTGSLTPPAVGLV 513

Db 503 RSGGCTDVNCEADPTICISCTCVNTAGSYTCECPD---FELNPTRVGCV 550

RESULT 14
ID 09WU9 PRELIMINARY; PRT; 2906 AA.
AC 09WU9:
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE FIBRILIN-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N.,
RA Wallner E.I., Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
morphogenesis of rat embryonic lung."
RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135060; AAD34439.1; -.
DR HSSP; P35555; 1EMN.
DR InterPro: IPR002086; Aldohyde_dehydr.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002212; TB.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS; PR000170; EGFBL00D.
DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2306 AA; 313372 MW; 9EB64E727044EF58 CRC64;

Query Match 12.9%; Score 375; DB 11; Length 2906;
Best Local Similarity 32.9%; Pred. No. 2,4e-21;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVNGGCEHACNAIPGAPRCQCPAGALADGRSCTASATQSCNDLCEHCVNPND- 301
DB 1237 EGMINGGCDPQCTNSESSEYCSSEGYALAMPDRSCA-----DIDE--CENNPDIC 1286
QY 302 -----QPGSSCMCEGTYRLAADQHRCEVDYDCILPSPCP-QRCVNTGGEGECHYR 353
DB 1287 DGGCTNIPGEYRCLCYGPFASMDKTCIDVNECDLNPNCMGEGECNTRGSGTCHQQL 1346
QY 354 NTDLVGE--CVEPVPDC--FRANCEYOCPLN-OTSILYCAEGFA-----PIR 398
DB 1347 GYSVKKATGCTD-VDEBEIGHNCDMAHSLNGLVSGFKSCRGSWVNGIKCIDLBDECA 1405
QY 399 HEPRHCQMFQNOTACPADCPNTQAS--CECEGYILDDGFTCTDIDE-----CENG- 448
DB 1406 NCTHCCSI-----NAOC-VNTPGSTRACASEGF-TGDGFTCSVDDECAENINLCENGQ 1456
QY 449 -----GFE-----GVCNMLPGFEFCICGPPDSALA 473
DB 1457 CLNVGATGCEGMEGFTASDSRSCQDIDECFQNTICVFQICNMLPGHFCICDDGGLD 1516
QY 474 RHIG--TDGD 481
DB 1517 RTGGHCTDID 1526

RESULT 15
ID 088840 PRELIMINARY; PRT; 3857 AA.
AC 088840;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE MUTANT FIBRILIN-1.
GN FBN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B10.D2;
RX MEDLINE-98069008; PubMed-9405934;
RA Bona C.A., Mural C., Casares S., Kasturi K., Nishimura H., Honjo T.,
RA Matsuda F.;
RT "Structure of the mutant fibrillin-1 gene in the tight skin (TSK)
mouse."
RL DNA Res. 4:267-271(1997).
DR EMBL; AF007248; AAC62317.1; -.
DR HSSP; P35555; 1ABV.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR002557; Chitin_binding.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR InterPro: IPR000822; Znf-C2H2.
DR Pfam: PF00008; EGF; 64.
DR Pfam: PF00683; TB; 12.
DR SMART; SM00494; ChIPD2; 2.
DR SMART; SM00179; EGF_CA; 60.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 61.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 50.
DR PROSITE; PS01187; EGF_CA; 61.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3857 AA; 418303 MW; 5BC061BEC527E04C CRC64;

Query Match 12.6%; Score 367; DB 11; Length 3857;
Best Local Similarity 30.3%; Pred. No. 1,4e-20;
Matches 118; Conservative 28; Mismatches 132; Indels 112; Gaps 21;

QY 212 SAAVAPLGL-----QIMCTAPPGAVOGHARAPGAMDCSVNGGCEHACNAIPGAPRC 265
DB 2148 SANLCFPHGRVNLIGKYQACNPGYHPTDRFLFCVIDIDECISINNGCEFTFNSDGSYEC 2207
QY 266 QCPAGALADGSGCTASATQSCNDLCEHCVNPND-----DPRGYSQMCETGYRL 316
DB 2208 SCQPGFALMPDRSCT-----DIDE--CEDNPICDGGCTNIPGEYRCLCYDGFA 2257
QY 317 AAQHRCEVDYDCILPSPCP-PQRCVNTGGEGECHYRPNYDLVGE--CVEPVPDC--FR 371
DB 2258 SEDMKTCVVDNECDLNPNCISGTCENTRGSGTCHQDMKYSKKGKGTGCTD-INCELEGA 2316
QY 372 ANCEYOCQPLNQT-STLYCAEGFA-----PIRPHRCQMFQNOTACPADCPN 420
DB 2317 HNDRAVAVNTAGSFKSCSPGMDIGIKYCTDIDECSSGTMCSON-----ADC-KN 2368
QY 421 TGAS--CECEGYILDDGFTCTDIDEC-EN-----GGF----- 450
DB 2369 TMGSYRCLCKDGI-TGDGFTCTDIDECSENILNCGGQLNAPGGYRCRCMDGFPSPADG 2427
QY 451 -----CS-----GVCNMLPGFEFCICGPPDSALAHIG-----TDGDSGK 484
DB 2428 KACEDIDECSLPNICVFQICNMLPGHFCRCECEIGYELDRSGGCTDVNCEADPTICISGN 2487

QY 485 -VDGDSGSGEPPTPGSTLTPPAVGLV 513
! : : !
Db 2488 CVNTPGSYTCDCPPD---FELNPTRVGCV 2513

Search completed: May 8, 2002, 12:44:52
Job time: 353 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:40:00 ; Search time 29.56 Seconds
(without alignments)
1329.705 Million cell updates/sec

Title: US-09-509-994-2

Sequence: 2916
1 MGVLYVIGALALAGLGFAP.....PSPYRQSTLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2912	99.9	575	1	THHUB
2	1903	65.3	577	2	thrombomodulin pre
3	1095.5	37.6	356	2	thrombomodulin - b
4	414	14.2	1574	2	MEGF6 protein - ra
5	414	14.2	1620	2	hypothetical prote
6	374	12.8	2907	2	fibriillin-2 precu
7	368	12.6	2871	2	fibriillin I - bovl
8	367.5	12.6	1184	2	fibriillin-2 precu
9	366	12.6	2918	2	fibriillin-2 precu
10	364	12.5	3002	2	fibriillin-1 precu
11	362	12.4	2871	2	fibriillin-1 precu
12	357	12.2	741	2	hypothetical prote
13	357	12.2	1221	2	fibriillin-2 precu
14	332.5	11.4	1964	2	notch4 - mouse
15	331	11.4	589	2	fibriillin-1D precu
16	330.5	11.3	1712	2	masking protein pr
17	324.5	11.1	689	2	fibriillin, splice fo
18	324.5	11.1	712	2	fibriillin 1, splice
19	322.5	11.1	1394	2	transforming growt
20	321	11.0	2321	2	notch3 protein - h
21	314.5	10.8	3507	2	hypothetical prote
22	313	10.7	798	2	hypothetical prote
23	312.5	10.7	1820	2	latent transformin
24	311	10.7	601	2	fibriillin 1 precu
25	309	10.6	683	2	fibriillin 1 precu
26	308	10.6	705	2	fibriillin, splice fo
27	303.5	10.4	1251	2	latent transformin
28	297	10.2	2531	2	Notch-1 protein -

30	287	9.8	2471	2	A49128	cell-fate determin
31	285	9.8	2352	2	T30201	Notch homolog prot
32	284.5	9.8	2437	2	S42612	Transmembrane prot
33	280.5	9.6	2703	1	A24420	transmembrane pr
34	280	9.6	2318	2	S45306	notch 3 protein - tr
35	279.5	9.6	387	2	I38449	extracellular prot
36	278	9.5	2531	2	S18188	notch protein homo
37	277.5	9.5	511	2	T17298	hypothetical prote
38	276.5	9.5	493	2	JC5621	epidermal growth f
39	273	9.4	2555	2	A40043	notch protein homo
40	269.5	9.2	1217	1	ECMSMG	epidermal growth f
41	264	9.1	1203	2	A49175	epidermal growth f
42	263.5	9.0	2531	2	T11070	notch B protein -
43	256.5	8.8	2524	2	A35844	notch homolog - se
44	254.5	8.7	674	2	I55476	notch protein - Af
45	249	8.5	678	2	B48089	growth potentialin

ALIGNMENTS

RESULT 1

THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Degashiki, Y.; Maru

J. Biochem. 103:281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed

A:Reference number: A41442; MUID:86227901

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SH1>

A:Cross-references: DDBJ:D00210; NID:g220126; PIDN:BA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritz, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of th

A:Reference number: A28307; MUID:87317665

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Degashiki, Y.; Nishioke, J.; Maruyama, I.; Zushi, M.;

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on

A:Reference number: A29680; MUID:88004395

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SU2>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: Lung endothelium

A:Note: part of this sequence, including the amino end of the mature protein, were

R:Wen, D.; Diltman, W.A.; Ye, R.D.; Deaver, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization

A:Reference number: A27073; MUID:88024950

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <MEN>

A:Cross-references: GB:M15552; NID:g339656; PIDN:AA59508.1; PID:g339657

A:Experimental source: placenta

A:Note: parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Okuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell,

Biochem. J. 295, 131-140, 1993
 A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.
 A:Reference number: S38954; MUID:94029900
 A:Accession: S38954
 A:Molecule type: protein
 A:Residues: 475-491, 'X', 493-494 <GER>
 A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chori R.Melinger, D.P.; Komives, E.A.
 A:Submitted to the Brookhaven Protein Data Bank, September 1995
 A:Reference number: A67369; PDB:1ZAO
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue R.Melinger, A.; Mathews, I.T.
 A:Submitted to the Brookhaven Protein Data Bank, August 1994
 A:Reference number: A52804; PDB:1HUT
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442
 R.Hirshal, R.; Komives, E.A.; M., F.
 A:Submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A65583; PDB:1FGD
 A:Contents: annotation; conformation by (1)H-NMR, residues 427-444
 R.Hirshal, R.; Komives, E.A.; M., F.
 A:Protein Sci. 5, 195-203, 1996
 A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the A:Reference number: A58595; MUID:96276211
 A:Contents: annotation; conformation by (1)H-NMR
 C:Genetics:
 A:Gene: GDB:THBD
 A:Cross-references: GDB:119613; OMIM:188040
 A:Map position: 20p11.2-20p11.2
 A:Introns: #status absent
 C:Complex: homodimer, urinary form
 C:Function:
 A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activation
 A:Pathway: blood coagulation/moderation
 A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie C:Keywords: thrombin complexed with the membrane-bound form is subject to endocytosis
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; beta-hydroxyaspartate; beta-hydroxyaspartic acid; blood coag protein
 E:1-18/Domain: signal sequence #status predicted <SIG>
 E:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
 F:19-513/Domain: extracellular #status predicted <EXT>
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>
 F:24-167/Domain: C-type lectin homology <LCH>
 F:177-199/Region: PEST sequence
 F:201-233/Region: PEST sequence
 F:245-280/Domain: EGF homology <EG1>
 F:288-323/Domain: EGF homology <EG2>
 F:329-362/Domain: EGF homology <EG3>
 F:369-404/Domain: EGF homology <EG4>
 F:408-439/Domain: EGF homology <EG5>
 F:445-480/Domain: EGF homology <EG6>
 F:485-513/Region: PEST sequence
 F:517-539/Domain: transmembrane #status predicted <TMN>
 F:540-575/Domain: intracellular #status predicted <INT>
 F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:174,223,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:245-255,252-260,288-296,292-308,310-323,329-340,336-349,351-362,369-376,374-38 F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:342/Modified site: erythro-beta-hydroxyaspartate (Asp) #status experimental
 F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental

Query Match 99.9%; Score 2912; DB 1; Length 575;
 Best local similarity 99.8%; Pred. No. 3e-166;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVVYLGALAGIGPAPAPAPGSGQVHDCFLVPGPTFLNASQICDGRGHM 60
 DB 1 MGVVYLGALAGIGPAPAPAPGSGQVHDCFLVPGPTFLNASQICDGRGHM 60
 QY 61 TVRSSVAADYISLLNGDGVGRRRLWIGLQLPGCCGPKRLGLRFGQWVGDNNTSYS 120
 DB TVRSSVAADYISLLNGDGVGRRRLWIGLQLPGCCGPKRLGLRFGQWVGDNNTSYS 120

DB 61 TVRSSVAADYISLLNGDGVGRRRLWIGLQLPGCCGPKRLGLRFGQWVGDNNTSYS 120
 QY 121 RMRRLDNPALCPGRCVAVSAEATVSEPTMEQCEYKADGFLCEHFPATCPPLAV 180
 DB 121 RMRRLDNPALCPGRCVAVSAEATVSEPTMEQCEYKADGFLCEHFPATCPPLAV 180
 QY 181 EPGAAAAYSTYGTGPPFAAGADFGALPVGSSAAVAPVGLQMLCTAPGAVQGHMAREAP 240
 DB 181 EPGAAAAYSTYGTGPPFAAGADFGALPVGSSAAVAPVGLQMLCTAPGAVQGHMAREAP 240
 QY 241 GANDCVENGSGEHNCNMPGAPRCQCPAGAAIADGSCASASTOSCNLCGHPVNP 300
 DB 241 GANDCVENGSGEHNCNMPGAPRCQCPAGAAIADGSCASASTOSCNLCGHPVNP 300
 QY 301 DPGSYSCMCEGYRLADQHRCEVDDCILPEPSPQRCVNTGGEFCHCYPNYDLVDG 360
 DB 301 DPGSYSCMCEGYRLADQHRCEVDDCILPEPSPQRCVNTGGEFCHCYPNYDLVDG 360
 QY 361 ECEPVPDPCFRANCEYQCPPLNQTSTYLCVCAEGFAPLPHEPHRQMCNCTACPAACDN 420
 DB 361 ECEPVPDPCFRANCEYQCPPLNQTSTYLCVCAEGFAPLPHEPHRQMCNCTACPAACDN 420
 QY 421 TQASCCEPGYILDDGFICTDIDECENGFCGVCVCHMLPGTFECICGSDALRHITGDC 480
 DB 421 TQASCCEPGYILDDGFICTDIDECENGFCGVCVCHMLPGTFECICGSDALRHITGDC 480
 QY 481 DSGKVDGDSGSGEPSPPTGTLTPPAGLVHSG 516
 DB 481 DSGKVDGDSGSGEPSPPTGTLTPPAGLVHSG 516

RESULT 2
 A60501
 thrombomodulin precursor - mouse
 N:Alternate names: fetomoduulin
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
 C:Accession: S08488; A32001; A60501
 R:Dittman, W.A.; Majerus, P.W.
 Nucleic Acids Res. 17, 802, 1989
 A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted A:Reference number: S08488; MUID:89128454
 A:Accession: S08488
 A:Molecule type: mRNA
 A:Residues: 1-577 <DIT>
 A:Cross-references: EMBL:X14432; NID:954781; PIDN:CA32597.1; PID:954782
 R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.
 J. Biol. Chem. 263, 15815-15822, 1988
 A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat A:Reference number: A32001; MUID:89008498
 A:Accession: A32001
 A:Molecule type: mRNA
 A:Residues: 97-577 <DIT>
 A:Cross-references: GB:J04060
 R:Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.
 Dev. Biol. 140, 113-122, 1990
 A:Title: Identification of fetomoduulin, a surface marker protein of fetal development A:Reference number: A60501; MUID:90292331
 A:Accession: A60501
 A:Molecule type: protein
 A:Residues: 19-22,330-343,479-489,545-555,562-575 <IMA>
 C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor
 F:24-165/Domain: C-type lectin homology <LCH>
 F:244-279/Domain: EGF homology <EG1>
 F:287-322/Domain: EGF homology <EG2>
 F:328-361/Domain: EGF homology <EG3>
 F:368-403/Domain: EGF homology <EG4>
 F:407-438/Domain: EGF homology <EG5>
 F:444-479/Domain: EGF homology <EG6>

Wed May 8 16:19:32 2002

us-09-509-994-2.rpr

Page 3

Query Match	65.38;	Score 1903;	DB 2;	Length 577;
Best Local Similarity	65.6%;	Pred. No. 2.9e-106;		
Matches 342;	Conservative 46;	Mismatches 125;	Indels 8;	Gaps 4;

Qy	1	MIGVIVLGLMLAGLGGPAPAEPEQPGSGSCVEHDEEALYPGAPAFPLNLSICGSLGKLM	60
Db	1	MIGVIVLGLMLAGLGGPAPAEPEQPGSGSCVEHDEEALYPGAPAFPLNLSICGSLGKLM	60
Qy	1	MIGVIVLGLMLAGLGGPAPAEPEQPGSGSCVEHDEEALYPGAPAFPLNLSICGSLGKLM	60
Db	1	MIGVIVLGLMLAGLGGPAPAEPEQPGSGSCVEHDEEALYPGAPAFPLNLSICGSLGKLM	60
Qy	61	TVRSSAADVYISLLISLNGSP--GVGRRRLAMIGLOLPECCGPKRLSGPLGFMWYGDNNSTX	119
Db	61	TVRSSAADVYISLLISLNGSP--GVGRRRLAMIGLOLPECCGPKRLSGPLGFMWYGDNNSTX	119
Qy	120	SRRAKRLDNGAPLRCGLPLCAVSAAEAVYBSEPIEMEQCEYKAGSGELCEHFPATRCPLA	179
Db	118	SRRAKRLDNGAPLRCGLPLCAVSAAEAVYBSEPIEMEQCEYKAGSGELCEHFPATRCPLA	177
Qy	180	VEP-GAAAAAAYSTGYTFPAARGADEQALPVGSSAAVADPLQLQMLCTAPRGAVOGHWABE	238
Db	178	VNTRDEPAHHISTYNTPEFGVSAGDEFQTLTPVGSSAAVEPLGLELTCRAPRPTSGHMAWE	237
Qy	239	APGAMPCSYENSGCECHACNATPGAPRCOCAGCAALOADRSGCTASATSCNDLCEHRCVP	298
Db	238	ATGAMNCSYENSGCEYLCTNRSTNEPRLCTPRMDLQADGRSCARPVYOSCEHCEHCVAS	297
Qy	299	NRPQPSYSYSCMEYSTRLAADHNCCEYDNCILPEBPCRCRVYVNOGGFECHCYPANPLY	358
Db	298	NAEVPYSYSCMEYSTRLAADHNCCEYDNCILPEBPCRCRVYVNOGGFECHCYPANPLY	357
Qy	359	DGECVEPVDPCFRANCEYOCQPLNQTSTLYCVAEBGPAFLPBEBRHQCMFCQTRCAPCDD	418
Db	358	DGECVLELDPGCSNCEQCOQVSPSTYRQICAPGPAFPRBDBPACHECFMEKNETSCPADCO	417
Qy	419	PNTQASCECPREYIILDDFTCTDIDECENMGFCISGVCHMLPTPEFICIGDPDASLARHIGT	478
Db	418	PNSPTVCECPREFIILDDSVCTDIDECESQGECTSCRNFPESYECICIGDPDALAQISK	477
Qy	479	DCDSGRV---DGGDGSGESEPPSPPTGSLTLPYAVLVYSG	516
Db	478	DCDPIVREDTKEEBGSSEPPSPPTGSLTLPYAVLVYSG	518

```

RESULT      3
A25918
thrombomodulin - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_charge 16-Jul-1999
C:Accession: A25918
R:Jackman, R.W.; Beeler, D.L.; Vandewater, L.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 83, 8634-8638, 1986
A:Title: Characterization of a thrombomodulin cDNA reveals structural similarity to the
A:Reference number: A25918; MVID:87067408
A:Accession: A25918
A:Molecule type: mRNA
A:Residues: 1-356 <TAC>
A:Cross-references: GB:114657; NID:9163762; PIDN:AA30785.1; PID:9163763
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: transmembrane protein
F:21-56/Domain: EGF homology <EG3>
F:64-97/Domain: EGF homology <EG3>
F:103-116/Domain: EGF homology <EG3>
F:143-178/Domain: EGF homology <EG4>
F:182-213/Domain: EGF homology <EG5>
F:219-253/Domain: EGF homology <EG6>

```

Query Match 37.6%; Score 1095.5; DB 2; Length 356;
Best Local Similarity 63.4%; Pred. No. 2e-58;
Matches 189; Conservative 31; Mismatches 63; Indels 15; Gaps 6

QY 229 CAVCGHWAREAPGAMDCSVENGSCERACNAITGATPRCCCPKAGALQADGRSCTASATOSC 288
 |||||
NH 5 GETEENRSREPPGMAGCVERGGCOHECKSGSAASTNCICPDPDALIQADGRSCGIPRAHPHC 64
 |||||

[illegible]

RESULT 4
T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like mo
A:Reference number: Z14126; MUID:98360089
A:Accession: T13954
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NAK>
A:Cross-references: EMBL:AB011532; NID:G3449293; PTDN:BAA32462.1; PTD:G3449294
A:Experimental source: strain Sprague-Dawley; brain
C:Genetics:
A:Gene: MEGF6

Query Match	14.28;	Score 414;	DB 2;	Length 1574;
Best Local Similarity	34.18;	Pred. No. 2.2e-17;		
Matches 104;	Conservative 28;	Mismatches 101;	Indels 72;	Gaps 17;

QY	224	CTAPPG-ANOGHMAEADGAMDCSEVNGGSEHNACNIPAPRCQCPACALADADGSRCTA	282
		1 1	
Db	147	CSRCPGQYADGD-GATCQDDYDCEAHNNGGQHNRCVMTPESTLYCECKPGFRLHTDGRCTL-	204
QY	283	SATQSC--NDLSEHCPV-----	309
		1 1	
Db	205	-ALSSCTYLAGGGGQHCQCTVLTQHCQCPROYOLOEDGRGVRKSPRABENGGCMHTCO	263
QY	310	-----CETGYRLADONHRCEDVDVDCILPESPQRQCVMTDGGFCGHCYPNVDL-VD	359
		1 1	
Db	264	ELRSLAHGCGHNPYQDLADRKTCSEDDYBECALGLAOCNAHGLCLMTPQSGFPCVCHAGYEGLAD	323
QY	360	G-EC-----VERPDPGFRAN-CETQOCQPLMOTSTLYCAGAGAPLPRPHEHNCQMF----	408
		1 1	
Db	324	GROCYRLIEMTIVSCAGNGGSHSCSH-TSTGPICTCPRGY-----ELDEDQRTCTCIDID	377
QY	409	---NOTACPADCDPNTQA--SCSECEPYIIL-DFGLCTDIDCEGNG-GEFCGVCYHNLPGT	461
		1 1	
Db	378	DCANSPPCQDAC-ANTPGYECSCFAGYRLMTDQCGCEDVDKACAGSHGCGEHNHNSNLGAS	436
QY	462	FECTIC 466	
		1 1	
Db	437	FQCF 441	

RESULT 5
T27283
hypothetical protein Y64G10a.f - *Caenorhabditis elegans*
CISpecies: *Caenorhabditis elegans*
CDate: 15-Oct-1999 #sequence_revision 15-Oct-1999
CAccession: T27283
R:Almscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: 220356
A:Accession: Y27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-References: EMBL:AL110498; NID:el542303; PIDN:CAB54471.1; CESP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CESP:Y64G10A.f
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7

Query Match 14.2%; Score 414; DB 2; Length 1620;
Best Local Similarity 27.0%; Pred. No. 2.2e-17;
Matches 128; Conservative 38; Mismatches 176; Indels 132; Gaps 20;

QY 119 YSHMARLDNGAPLGGPCVAVSAE-----ATPSPRMEVQCE--YVADGFLCEE 169
DB 56 YLRFAFRSRRGSKCCLLRQANGSADLCHNGTCVPSHNDNEQVCEPVGFTAKQY 115
QY 170 HEPATCRPLAVPPGAAAASITTYGTFPAARGADFOALPYGSSAAVAPLGLQIMCTAPPG 229
DB 116 D-ANEC--MANNGCEHCVN-TIGTY-----CRCMPG 145
QY 230 AVQGHWAREAPGAMDCSYENGCEHCACNAIPGAPRCQCPAGALQADGRSTASATQSC- 288
DB 146 FELSGDNATCSDIDCAVANGCSDRCVNSPGFRCDCPSDLYLHADGRTG--GKVTSCS 203
QY 289 --NDLCEHCVPNPDPQPSYSCMCEGYRLAADHRCCEVDCLLEPSPRCVNTNNGG 346
DB 204 TDNGGCEHC--ENDSNGEFTYRCRVRGFKLSEKRSQVDFPDNKGCGCHHCKNNHGR 262
QY 347 FECHCYPNYDL-VDECEVPEYDPCFRAN-CEYOCQPLNOTSYLCVACGAFAPRPHPRRC 404
DB 263 ACCOCYCPGPHLSYDRSCVDIDECANNGCEHCENVKGT-YXCKCREGY-OLGSDGRTG 320
QY 405 OMF-----CNOTACPADC--DPNTQASCEPEGYTL-----433
DB 321 EEMLGCCQYNGGCGHDCYDQPDGHHCKCRNGYILANQKICHONISYVIAHARPLMD 380
QY 434 -----DGGF-----ICTDIDEC-ENGCGCGVC 455
DB 381 STEYTCVTPYDLTCHKLCMHLDGSHVQCFCDGDELIDSKRFQDINCEHNNGCGSIC 440
QY 456 HMLPTFEFCICGPDALAHRTG-----DCDS-----GKVDGSGSGSEP 495
DB 441 VNLASGVCEQCKPFRMLKDRKTCEDISECSSNNGCGCEOTCSNOEGYWCSEP 494

RESULT 6
A57278
Fibrillin-2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Feb-1996 #sequence, revision 23-Feb-1996 #text_change 11-Jan-2000
C:Accession: A57278
R:Zhang, H.; Hu, W.; Ramirez, F.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix
A:Reference number: A57278; MUID:95263670
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-References: GB:139790; NID:9762830; PIDN:AAA74908.1; PID:9762831
F:1233-1274/Domain: EGF homology <EGF1>
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.8%; Score 374; DB 2; Length 2907;
Best Local Similarity 32.9%; Pred. No. 8.3e-15;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVENGCCEHCACNAIPGAPRCQCPAGALQADGRSTASATQSCNDLCEHFCVPPND-- 301
DB 1238 ECMIMNGCCTDQCTNSGSEYSCSEGVYALMDGSRCA-----DIDE--CENNPDIC 1287
QY 302 -----QPSYSCMCEGYRLAADHRCEDVDCLLEPSPC--QRCVNTNNGGCEHCY 353
DB 1268 DGGCCTNPGEYKRCIDYDGFMAEMDKITCVDNECDLNPICLGTENTKSGFICHDM 1347
QY 354 NYDLVDE--CPEVPDPC--FRANCEYQCPLN-OTSLCYCAGGFA-----PIP 398
DB 1348 GYSVAKGTGTGTD-VDECEIGAHNCMDHASCINVPSEFSCSRSRGVWNGIKICIDIDDECA 1406
QY 399 HEPHRCQMFNCNOTACPADCDPNTQAS--CEPEGYILDDGFTCTDIDE-----CENG- 448
DB 1407 NGTHQCSF-----NAGC-VNTPGSRACSEGF--TGDGFTCSVDDECAENNLCENQ 1457
QY 449 -----GF-----GCHNLPGTFEFCICGPDALAH 473
DB 1458 CLNPGARCECEMGFTPASDSRSCODIDECFQNTICVFTCNMLPGMFICIDGDEYLD 1517
QY 474 RHIG--TDCCD 481
DB 1518 RTGGNCTDID 1527

RESULT 7
A55567
Fibrillin 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence, revision 20-Feb-1995 #text_change 11-Jan-2000
C:Accession: A55567
R:Willstra, D.; J.; L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization
A:Reference number: A55567; MUID:95137597
A:Accession: A55567
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <WIL>
A:Cross-References: GB:128748; NID:9508427; PIDN:AAA74122.1; PID:9508428
F:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.6%; Score 368; DB 2; Length 2871;
Best Local Similarity 31.5%; Pred. No. 1.9e-14;
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 244 DCSVENGCCEHCACNAIPGAPRCQCPAGALQADGRSTASATQSCNDLCEHFCVPPND-- 300
DB 1200 ECTIMNGCCEFTNSGSEYSCSEGVYALMDGSRCA-----DIDE--CENNPDIC 1249
QY 301 -----DQPSYSCMCEGYRLAADHRCEDVDCLLEPSPC--QRCVNTNNGGCEHCY 353
DB 1250 DGGCCTNPGEYKRCIDYDGFMAEMDKITCVDNECDLNPICLGTENTKSGFICHDM 1309
QY 354 NYDLVDE--CPEVPDPC--FRANCEYQCPLN-OTSLCYCAGGFA-----PIP 398
DB 1310 GYSVAKGTGTGTD-VDECEIGAHNCMDHASCINVPSEFSCSRSRGVWNGIKICIDIDDECS 1368
QY 399 HEPHRCQMFNCNOTACPADCDPNTQAS--CEPEGYILDDGFTCTDIDE-----EN- 447
DB 1369 NGTHQCSF-----ADC-KNTMGSRACSEGF--TGDGFTCTDIDDECAENNLCENQ 1419
QY 448 -----GF-----GCHNLPGTFEFCICGPDALAH 473
DB 1420 CLNPGARCECEMGFTPASDSRSCODIDECFQNTICVFTCNMLPGMFICIDGDEYLD 1479
QY 474 RHIG-----TDCCSGK-VDGGSGSGSEPPSPPTPGSTLTPPAGLV 513
DB 1480 RSGGCTDVECTDPTTCTISGNCVNTPGSYTCDPDD---FELNTPRVGCV 1527

RESULT 8
A55184
Fibulin-2 precursor - human
N:Alternate names: protein DKFZP586A1519.1
C:Species: Homo sapiens (man)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: A55184; MUID:95104855
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
Genomics 22, 425-430, 1994
A>Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene
A:Reference number: A55184; MUID:95104855
A:Accession: A55184
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1184 <ZHA>
A:Cross-references: GB:X82494; NID:9575232; PID:9575233
R:Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08744
A:Molecule type: mRNA
A:Residues: 656-719, ODECLMGAHDCSRROCFVNTLGSFYCVNHTVLCADGYILNAHRKCYD, 720-853, 'T', 855-1114
A:Cross-references: EMBL:AL050095
A:Experimental source: adult uterus; clone DKFZP586A1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135821
A:Map position: 3p25-3p24
A>Note: DKFZP586A1519.1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; extracellular matrix
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>
F:905-941/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367.5; DB 2; Length 1184;
Best Local Similarity 30.0%; Pred. No. 1e-14;
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

154 EEOQCEVAD-GFLCEFFPATC-----RPLAV-----EGGAAAVSTTGPFAARG 201
DB 531 EGOSESNPNLGPCLN-HVMTSCGEBPLVPEYRRPPEAAARRYS-----EEM 582
QY 202 ADFQALPGSSAAV---APLGLQIMCTAPPGAYOGHMAKRAFGAMDSCVENGCEHACNA 258
DB 563 AGREALSLGTLELRLNSLPBGDDQDECLLPGEL-----COHLICIN 622
QY 259 IPGAPRCOPAGALQADGRSC-----TASAT-----QS 287
DB 623 TVGSYHACAFPGFSLQDDGRTRCPRGHPPEAPQEPALKESEFVASNTIPLPLPONT 682
QY 288 CND---LCEHFCYRPNDDGSGYSCKMETGYRLAADQHRDEDVDDCILESPCP--QRQVNT 343
DB 683 CNDNGPCQKQVC---STVGSAICSCFPGYATIMADGVSEDIINECVTDLHTSCRGHCYNT 739
QY 344 OGGEFCH---CYPNYDVLDECEVPVPCFRANCEYOCOPILNOSTYLCV-----389
DB 740 LGSFHCYKALCEPGLYALKDEBC-EDVDEC--AMGTHICQF---GLFCQTKTSFYCOA 792
QY 390 ---CAEGFAPLPH-----EPHRCQMFQNTCAPADCPNTQASCECPGY-I 432
DB 793 RORCMGELQDPREGNCVDINECTLSLSEPCRPFGSCINTVGSYTCORNLFI---CARGYHA 849
QY 433 LDDGFLCTDIDECENGFCGSG---VCHNLPGTFECTIC 466
DB 850 SDDGAKCYDVNDECTGVHRCGEGVCHNLPGSYRCDC 886

C:Species: Homo sapiens (man)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 17-Nov-2000
C:Accession: A54105; S17063; S31101
R:Zhang, H.; Apfelfroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecha
J. Cell Biol. 124, 855-863, 1994
A>Title: Structure and expression of fibrillin-2, a novel microfibrillar component pr
A:Reference number: A54105; MUID:94165150
A:Accession: A54105
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-2918 <ZHA>
A:Cross-references: GB:U032272
R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
Nature 352, 330-334, 1991
A>Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
A:Reference number: S17063; MUID:91304567
A:Accession: S17063
A:Molecule type: mRNA
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>
A:Cross-references: EMBL:X62009
R:Maliewicz, D.M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S31101
A:Accession: S31101
A:Molecule type: mRNA
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P',
A:Cross-references: EMBL:X62009
C:Genetics:
A:Gene: GDB:FBN2
A:Cross-references: GDB:128122; OMIM:121050
A:Map position: 5q23-5q31
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: extracellular protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-2918/Product: fibrillin-2 #status predicted <MAT>
F:1245-1280/Domain: EGF homology <EGF1>
F:1970-2013/Domain: EGF homology <EGF2>

Query Match 12.6%; Score 366; DB 2; Length 2918;
Best Local Similarity 33.1%; Pred. No. 2.5e-14;
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

QY 244 DCSVENGCEHACNAIPGAPRCOPAGALQADGRSCASATQSCNDLCEHFCVNPDP-- 301
DB 1244 ECMIMNGCGDTQGTNEGSGECSGVALMPDGRCA-----DIDE--CENNPDC 1293.
QY 302 -----QPGSYSCMETGYRLAADQHRCEEDVDDCILESPCP--QRQVNTQGFECYCP 353
DB 1294 DGGQCTNIPGELYKCLCYDGFMAKMDKKTIDVNECDLNSNICMFGCEYTKGSFICHQL 1353
QY 354 NYDLVGE--CYEVPVPC--FRANCEYOCQPLN-QTSYLCVCAEGFAPLPH-----EPHR 403
DB 1354 GYSYKKTGTCTD-VDECELGANHCMDHMSCLNIPGSFSCSCEHG--IGNGIKCIDLDE 1410
QY 404 CQFNCQNTAPADCPDPTQAS--CECPGYILDDGFICTDIDE-----CENG-----448
DB 1411 CSNGTHQCSINAGC-VNTPGSYRCACSEGF--TGDGFTCSGDVDECAENMLCENGQCLANVP 1468
QY 449 -----GF-----CSGVCHNLPGTFECTICGPDNALRHIG- 477
DB 1469 GAYRCCEMGFTPADSPSRSCQDIDECSPQNICVSGTCNNLPGMFHCICDDGYELDRGTGN 1528
QY 478 -TDGD 481
DB 1529 CTDID 1533

RESULT 10
A47221
Fibrillin 1 precursor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jun-1995 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chaberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perella, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglman, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 132-3002 <PER>
 A:Cross-references: GB:U3923; NID:9306745; PIDN:AB02036.1; PID:9306746
 R:Maalen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556
 R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831
 A:Accession: I59574
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2217-2288, 'T', 2290-2325 <RES>
 A:Cross-references: GB:S54426; NID:9264860; PIDN:AB25244.1; PID:9264861
 R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.
 Nature 352, 330-334, 1991
 A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe
 A:Reference number: S17062; MUID:91304567
 A:Accession: S17062
 A:Molecule type: mRNA
 A:Residues: 'VUVTYVFTLSYKML', 944-1444 <LEE>
 A:Cross-references: EMBL:X62008; NID:931398; PIDN:CA55534.1; PID:95924015
 A:Accession: S62111
 A:Molecule type: Protein
 A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE>
 R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21381-21385, 1989
 A:Title: Connective tissue microfibrils. Isolation and characterization of three large F
 A:Reference number: A34198; MUID:90078246
 A:Accession: A34198
 A:Molecule type: Protein
 A:Residues: 565-575, 1890-1892, 'T', 1894-1900 <MAD>
 C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
 C:Genetics:
 A:Gene: GDB:FBNI
 A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
 A:Map position: 15q21.1-15q21.1
 A:Introns: 2236/1; 2258/1; 2297/1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M
 F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted
 F:1332-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MAD>
 F:1332-1367/Domain: EGF homology <EGF2>
 F:1457-1492/Domain: EGF homology <EGF1>
 F:2262-2295/Domain: EGF homology <EGF1>

Query Match 12.5%; Score 364; DB 2; Length 3002;
 Best Local Similarity 30.8%; Pred. No. 3.4e-14;
 Matches 11; Conservative 22; Mismatches 105; Indels 122; Gaps 20;

QY 244 DCSVNGGCEHACNAIFGAPRCOCAGALQADGNSCTASATOSCNDCLEHFCVNP--- 300

Db 1331 ECSTIMNGCTFCTNSGSESCQPFALMPQRCST-----DIDE--CEDPNIC 1380
 301 -----DQPSYSCMETGYRLAADHCEVDVDCILPSPC-PORCVNTGGCFECHYR 353
 Db 1381 DGGCTNIPGEYRCLCYDGMASEDMKTCYDVNECDLNNICLSGTCENTKSFICHO 1440
 354 NYDLVDGE-CVEPVDP-C-FRANCYQCOPLNQT-SYLCVCAEGFA-----PIP 398
 Db 1441 GYSOKKKTGTCTD-INCEIGAHNCGHACVTNAGSFKSCSPGWIIGIKCTDDECS 1499
 399 HEPRHRCMFCNQACAPDCDPNQTAS--CEPCGYTLLDGFCTDIDE--EN----- 447
 Db 1500 NGTHMCSQH-----ADC-KNTMGSTRCLCKEGY-TGSGFTCTDIDESENILCNGQ 1550
 448 -----GGF-----CS-----GVCHNLPGTEFCICGPDSDALA 473
 Db 1551 CLNAPGTYRCEDMGFPYPSADGKACEDIDECSLPNICVGTGNNLGLRCEGEIYELD 1610
 474 RHIG-----TDCSGKVDGSGSGEPSPPTGCS-----TLTPAVGLV 513
 Db 1611 RSGGNCYDVNECDLPTCTIGSNCVN-----TPSGYICDCPPDPPELMPTRVGCV 1658

RESULT 11
 A53624
 Fibrillin-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 23-Mar-1995 #sequence_rev150n 23-Mar-1995 #text_change 11-Jan-2000
 C:Accession: A55624
 R:Yin, W.; Smiley, E.; Gemmiller, J.; Sangalnet, C.; Lawton, T.; Perella, L.; Ramire
 J. Biol. Chem. 270, 1798-1806, 1995
 A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin
 A:Reference number: A55624; MUID:95130561
 A:Accession: A55624
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 12871 <YIN>
 A:Cross-references: GB:U29454; NID:9575509; PIDN:AAA56840.1; PID:9575510
 C:Genetics:
 A:Gene: Fbn-1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.4%; Score 362; DB 2; Length 2871;
 Best Local Similarity 29.4%; Pred. No. 4.3e-14;
 Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;

QY 212 SAAYAPVIGL-----QLMCTAPRGAVGHNARARAPGAMQCSYVNGGCEHACNAIFGAPRC 265
 1162 SANLCPPHRCVNLIGKYQACNPYHPTHDLFCVDIDECSIMNGCETFTNSDGYEC 1221
 266 QCPAGALQADGRCSTASATOSCNDCLEHFCVNP-----DQPSYSCMETGYRL 316
 Db 1222 SCQGFALMPDGRST-----DIDQ--CEDPNICDGGCTNIPGEYRCLCYGFA 1271
 317 AADQHRCEVDVDCILPSPC-PORCVNTGGCFECHYRNYDLVGE-CVEPVDP-C-FR 371
 1272 SEDMKTCYDVNECDLNNICLSGTCENTKSFICHOEMGSGKRGKGTCTD-INCEIGA 1330
 QY 372 ANCYQCOPLNQT-SYLCVCAEGFA-----PIPHPRHCQMFQNTACAPACDPN 420
 1331 HNCGRHACVTNAGSFKSCSPGWIIGIKCTDIDECSNGTHMCSQH-----ADC-KN 1382
 421 TQAS--CEPCGYTLLDGFCTDIDE--EN-----GGF----- 450
 Db 1383 TMGSTRCLCKDGY-TGSGFTCTDIDESENILCNGQCLNAPGTYRCEDMGFPYPSADG 1441
 451 -----CS-----GVCHNLPGTEFCICGPDSDALRHIG-----TDCDSGK 484
 Db 1442 KACEDIDECSLPNICVGTGNNLGLRCEGEIYELDRSGGNCYDVNECDLPTCTIGSN 1501

QY 485 VDGSDSGSGPPSPRPS-----TLTPPAGLV 513
 DB 1502 CVN-----TPGSTYCCSCSPDELPNTRVGCV 1527

RESULT 12

46488
 hypothetical protein DKFZp34J065.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46488

R:Diesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035

A:Accession: T46488

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-741 <AAA>

A:Cross-references: EMBL:AL137638

A:Experimental source: adult testis; clone DKFZp34J065

A:Note: DKFZp34J065.1

Query Match 12.2%; Score 357; DB 2; Length 741;
 Best Local Similarity 31.6%; Pred. No. 2, 9e-14;
 Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;

QY 245 CSVENGCENACNAIPGAPRCQCPAGALQADGRSCTA-SATQSCNDLCENFCVPNPDP 303
 DB 68 CAMEDHNCBOLCVNVPSCYCCSGYALAEDEKRCVAADYCASEHNGCEHCY-NAD-- 124
 QY 304 GSYSCMCETGYRLAADHRCEDVDCTLEPSPCRPCVNTQGGFECCHCYNNYL-VDGR 362
 DB 125 GSYLCQCHGFAALNPDKCTKIDYCASSNHGCONHECVNTDSTSCHLKGFTLNPDKKT 184
 QY 363 VEPYPCF--RANCYQCPINOTSYLCVCAEGFAPIPH-----EPHRCQMF 408
 DB 185 CRRINCYALKPCGECHECVNME-SYCCRCHRGVTLDPNCKTCSRVDCAQDQHGCEQ 243
 QY 409 NQTCAPADCPNQA--CECPREYITDDGF-ICTDIDEC--ENGFGCGVCHNLPGRF 462
 DB 244 -----LNTEDSFVCCSCSEFLINEDLKTCSRVDYCLSDHG--CEYSCVNDRSP 291
 QY 463 ECLGPDALARIHTDCDSGKYDG--GDSC 491
 DB 292 ACQC-PGHHVLRSDGKTC--AKIDSCALSDHG 320

RESULT 13

A49457
 fibulin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000

C:Accession: A49457; S74095

R:Pen, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.
 J. Cell Biol. 123, 1269-1277, 1993

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with
 A:Reference number: A49457; MID:94064787

A:Accession: A49457

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1221 <PAN>

A:Cross-references: GB:X75285; MID:9437046; PIDN:CA53040.1; PID:9437047

R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996

A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met

A:Reference number: S74094; MID:96439073

A:Accession: S74095

A:Molecule type: protein

A:Residues: 226-238, 'X', '240-247:260-275;336-344, 'L', '346-361;405-426;566-568, 'EW', '569-589

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer

F:942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;
 Best Local Similarity 29.5%; Pred. No. 4, 3e-14;
 Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

QY 154 EQQCEKAD-GEICFEFFPAPC-----RPLAV-----EPGAAAAS-----IT 192
 DB 521 EQQSCENPNILGYPCN-HVMSLCEGEPPLIVPEVRRPEPEAPARRVSEMEMASREALS 579
 QY 193 YET-----PFAARGAPFOQ--ALP-----VGSSAAAPGLGLM-----CTAPRG 229
 DB 580 LGTEALEPNSLPDGDQDECLMLPGELCONLCTNVTGSTRCACPGPELGGDGRCPDRG 639
 QY 230 AVQGHMARE-APGAWDCSV-----ENGCEHACNAIPGAPRCQCPAGAA 272
 DB 640 APQLDFARESAPRSASQVSPNTPPLVPYQPNCKDNPCQVGRVYGDFTAMCSCFFGYA 699
 QY 273 LDADRSC-----TASATQSCNDLCENFCVPNPDPQPGSYSC-----MCEGYRLAADQHR 322
 DB 700 IMADVSCDEDDDECLMGTHDCS--WKQPCV--NTLGSFYCVNHTVLCAGEYILNA-HRK 753
 QY 323 CEDVDCTLEPSPC--PQRCVNTQGGFECCH--CYPNYDLVDGCVEPVDPGCFRA--NC 374
 DB 754 CYVDINECTYDLHTCTRAHCYNTPGSFQCYALNCEPGYVLTDECTD-VDECVTGHNC 812
 QY 375 E--YQCPINOTSYLCV--CAEGFAPIPH-----EPHRCQMFQNOTCPA 415
 DB 813 QAGFCQN-TKGSFYQARQCMDFLODPGNCVINDICTSLPEPSCGSCINTVGSY 871
 QY 416 DCDPNTQASCCEPREG-ITDGFICTDIDECENGFGCG--VGHNLPGEFCICGR 468
 DB 872 TCQRNPLV--CGRTYHANEBSCEVDVNECTGVHRCGSQLCTYNLPGSTRCDCKP 925

RESULT 14

T09059
 notch4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jan-2000

C:Accession: T09059

R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Loretz, C.;
 submitted to the EMBL Data Library, October 1997

A:Description: Sequence of the mouse major histocompatibility locus class III region.

A:Reference number: 216543

A:Accession: T09059

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1964 <ROW>

A:Cross-references: EMBL:AF030001; MID:92564945; PID:92564947

C:Genetics:

A:Gene: notch4

A:Map position: 17

A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1;
 1679/3; 1729/1; 1761/3

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: receptor; signal transduction

F:514-545/Domain: EGF homology <EGF>

Query Match 11.4%; Score 332.5; DB 2; Length 1964;
 Best Local Similarity 26.3%; Pred. No. 1, 8e-12;
 Matches 128; Conservative 34; Mismatches 142; Indels 183; Gaps 31;

QY 134 GFLCAVAASAEITVSEPIWEQCEVADDFICE-FHFPTGCPRLAVPECAAAVSTIT 192
 DB 35 GGTCTLRSLRGQIC-----QC--APGLDETQCFPPPCR----- 66
 QY 193 YETPRAAGADFOAL--PVGSSAAVAPLGLQICTAPPG-----AVQGHMAREAPGAWDC 245
 DB 67 -DTQCKNGSGQALLPFPSPRSRSTSLTHTPSCCTCPSGTGRGCTHLELCPSPF-C 124

QY 246 SVENGGCHACNAIPGAPRCQCPAGAAIQADGRSCTASATQSCNDLCEHFCVNPDPG- 304
 Db 125 S--NGG--HCYVQASGRPQSCSEPGMT-----GEQCQ-----LRDCSANPCAMNG 166
 QY 305 -----SYSCMCEGRLADQGRCE-DVDDCLLESPCRQ--RCVNTQSGFECHC--- 351
 Db 167 VCLATVPQIQRCPPGR---EGHTECEHDINECFLEPGRCPQGTSCNTLGSQCLCPVG 222
 QY 352 --YPNYDLVDECEVPPVPCFRANCEYOCQPL---NOTSYLCVCAEGFAPIPHE----- 400
 Db 223 QEGPQCKRKAC--PPESCLNG--TCQLVPEGHSTPHLCCLPPGFGLDDEMNDDC 277
 QY 401 -PHRCOMFCNOTACPADCDPNTQASCECPREYIIDDGFICT-DIDCE-----NGGF 450
 Db 278 VRHOCQ---NGATCIDGLDITYT---CLCPKTV---KGMDCSEDIDECENAGPPRCRNGT 328
 QY 451 C-----SGVCHNLPGTFPCICGP----- 468
 Db 329 CQNTAGSFHCVCVSGWAGAGEENLDCAATCAPGSTCIDRVGSFSCCLCPRTGLLCH 388
 QY 469 --DSALAR--HIGTDCDSGKVDG-----GDSGS-----GEPPSP----- 499
 Db 389 LEDMCLSGPCHVNAQCSINPLTGSTLCIQPGYSGSTCHODLDECQMAQGPSPCEHGS 448
 QY 500 ---TPGS 503
 Db 449 CINTPGS 455

RESULT 15

T43210
 fibulin-1D precursor - Caenorhabditis elegans (fragment)
 C:Species: Caenorhabditis elegans
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
 C:Accession: T43210
 R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
 submitted to the EMBL Data Library, June 1998
 A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character
 A:Reference number: Z22337
 A:Accession: T43210
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-589 <BAR>
 A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1
 C:Genetics:
 A:Note: Intron positions not resolved (incomplete sequence)
 C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4%; Score 331; DB 2; Length 589;

Best Local Similarity 28.1%; Pred. NO. 8.4e-13; Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

QY 103 GPLRGF-QWVTGDNNTSYSRMARLDLNGAPLGPVAVSAEATVPSEPT-----WEEQ 156
 Db 19 GGLRSEFNKCCMGDLEITH---ASELITGRPLNDPHVHLIGDRGASSHCHEHCHDRGGEKV 75
 QY 157 OCEVKA-----DGLCFEHP--ATCRPLAVE-----PGAAAAVSTYGTFFA--- 198
 Db 76 ECSCRSRGLDAPDGMACVDHIDECATLMDCLIESQRLNTPGSFKIRTLSCGTGAMDS 135
 QY 199 --ARGADFOALPYGSSAVALPLGLQMLCTAPPAVO-----GHMAREAPGAMDCS--- 246
 Db 136 ETERCRDYDECNLGSH---DCGLPYOCRNTQGSTYRCDAKKCGDGELONPMTGECTSITC 191
 QY 247 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGMALQADGRSCTA 282
 Db 192 PNGYYPKNGMCMNDIDECVTGNHCGAGECVNTPGSFRQCGKGNLCAGYEYN----- 243
 QY 283 SATQSCNDL--CEH-----FCVNPDPQPGSYSCMCEGRLADQGRCEVDYDCIL--- 331
 Db 244 GATGCEDEVNRECQCGVCGSMCI---NLPGTYKCKCGPGYEFENDAKKRCEDEVDECIFAG 300

QY 332 EPSPCPQRCVNTQGGFECHCYPNYDLV--DGECEVPEVDCFR--ANCEYOCQPLNQTSLC 388
 Db 301 HVCDSLAEICINTIGSFECKKRGFQLASDGRCEDEVNCTGTGIAECQKCVNI--PGSYOC 359
 QY 389 VCAEGFAPIP-----HEPHROMEC--NOTACPADCDPNTQAS--CCCPREYIIDDGF 438
 Db 360 ICDRGFALGPDTGCKEDIDECSTIMAGSNDICMGCC--INTKGYLCCQPPGYKIQPDGRT 418
 QY 439 CTYDIDCEENGFCGSG---VCHNLPGTFPC 464
 Db 419 CVDVDECAMGE--CAGSDKVCVNTLGSFKC 446

Search completed: May 8, 2002, 12:40:05
 Job time: 251 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:38:52 ; Search time 42.86 Seconds
(without alignments)
891.783 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MGLVLTGALALAGLGFAP.....PSPTPGSTLTTPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2916	100.0	516	20	AAV09348 Human thrombomodul
2	2916	100.0	516	21	AAV83935 Human thrombomodul
3	2912	99.9	516	13	AAK23016 Truncated human th
4	2912	99.9	516	20	AAV09347 Human thrombomodul
5	2912	99.9	516	21	AAV83934 Human thrombomodul
6	2912	99.9	516	21	AAV69529 Human thrombomodul
7	2912	99.9	516	12	AAK11534 Human thrombomodul
8	2912	99.9	516	14	AAK41806 Human thrombomodul
9	2912	99.9	516	14	AAK43031 Human thrombomodul
10	2910	99.8	515	12	AAK10617 Soluble thrombomod
11	2910	99.8	515	9	AAK82070 Human thrombomodul

12	2908	99.7	516	13	AAK22018 Human thrombomodul
13	2908	99.7	516	13	AAK31572 Human thrombomodul
14	2904	99.6	516	14	AAK22017 Human thrombomodul
15	2903	99.6	516	13	AAK20639 Human urinary thro
16	2879	98.7	516	20	AAW73970 Human thrombomodul
17	2862	98.1	516	13	AAK22189 Sequence of thromb
18	2826	96.9	498	16	AAK84185 Human derived thro
19	2826	96.9	498	18	AAW01600 Human thrombomodul
20	2826	96.9	498	21	AAV67402 Novel sugar chain-
21	2824	96.8	497	17	AAK94607 Human recombinant
22	2820	96.7	497	19	AAW69520 rSTM protein SEQ I
23	2807	96.3	500	21	AAV69530 Human thrombomodul
24	2804	96.2	494	21	AAV67401 Novel sugar chain-
25	2766	94.9	494	16	AAK78726 Recombinant thromb
26	2764	94.8	494	13	AAK24400 Recombinant thromb
27	2749	94.3	494	16	AAK78727 Thrombomodulin wit
28	2744	94.1	486	12	AAK13877 Thrombin-binding s
29	2716	93.1	480	13	AAK22013 Truncated human th
30	2686	92.1	475	13	AAK22032 Mature thrombomod
31	2680	91.9	476	16	AAK78725 Modified thrombomo
32	2680	91.9	476	16	AAK86376 Modified thrombomo
33	2678	91.8	476	16	AAK86377 Sequence of human
34	2661	91.3	535	9	AAK80641 Thrombomodulin ana
35	2623	90.0	462	15	AAK45336 Thrombomodulin ana
36	2619	89.8	462	15	AAK45337 Thrombomodulin ana
37	2619	89.8	462	15	AAK45338 Thrombomodulin ana
38	2602	89.2	462	15	AAK45339 Thrombomodulin ana
39	2600	89.2	462	15	AAK45350 Thrombomodulin ana
40	2595	89.0	462	15	AAK45343 Thrombomodulin ana
41	2576	88.3	462	15	AAK45349 Thrombomodulin ana
42	2575	88.3	462	15	AAK45353 Thrombomodulin ana
43	2570	88.1	462	15	AAK45347 Thrombomodulin ana
44	2566.5	88.0	461	15	AAK45335 Thrombomodulin ana
45	2565	88.0	462	15	AAK45346 Thrombomodulin ana

ALIGNMENTS

RESULT 1
ID AAY09348 standard; Protein; 516 AA.
AC AAY09348:
DT 08-JUL-1999 (first entry)
XX Human thrombomodulin SEQ ID NO:2.
DE Human: thrombomodulin; aqueous parenteral solution; storage;
KW distribution; acute coronary syndrome; thrombosis; embolism;
diabetes.
XX Homo sapiens.
OS
XX
XX WO918994-A1.
XX
XX
XX 22-APR-1999.
XX
XX 13-OCT-1998; 98WO-JP04609.
XX
XX 11-NOV-1997; 97JP-0308523.
XX PR 15-OCT-1997; 97JP-0281659.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
XX WPI: 1999-277444/23.
XX N-PSDB; AAX55880.
XX Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,

PT thrombosis, embolism, and diabetes
XX
PS Claim 7; Page 90-92; 97pp; Japanese.
XX
CC The present invention describes a method for maintaining the quality of
CC an aqueous parenteral solution of thrombomodulin comprising buffer and
CC surfactant aseptically filled in a case or syringe. Maintaining the
CC quality of an aqueous, parenteral thrombomodulin solution is
CC characterised in that the solution: (a) comprises soluble thrombomodulin,
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
CC filled into: (1) a case or (11) a syringe without any empty space; and
CC (c) is kept in liquid form in storage and distributed and not frozen or
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome
CC (such as myocardial infarction, unstable angina and coronary artery
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
CC lupus erythematosus or Barrett's syndrome), multiple organ failure,
CC disseminated intravascular coagulation, transient ischaemia, diabetes,
CC liver veno-occlusive diseases and deep vein thrombosis. The composition
CC is stable for a long period of time and can be stored and distributed in
CC ready to use form avoiding the problems of dissolution and accuracy when
CC preparing on demand. The present sequence represents human
CC thrombomodulin.
XX
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVGALALAGLPAPAEPOPGSGQCEHDFALYPPGATFLNASQICDGLRGHLM 60
DB 1 mlgvlvgalalaglgfpapaepppgsgqcevdcfalypgpatflnaasqicdglrghlm 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPRRLGLRGFWVTGDNNTSYS 120
DB 61 tvrssvaadvysllngdgvgrrrlwigqlppgcgdprrlglrgfwvtdnntsys 120
QY 121 RMARLDLNGAPLPGPLCVANSAAEATVPSEPIWEBOQCEVKADGFLCEHFPATCRPLAY 180
DB 121 rmarldlngaplcgplcvansaeeatvpsepiweeqcevkadgflcehfpatcrplay 180
QY 181 EPGAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGIQLMCTAPPGAVGCHMARBP 240
DB 181 epgaaaavsitlygtpfaargadfgalpyvssaaavaplgqlmctappgavghmarbp 240
QY 241 GAWDCSVENGCEHACNAIPGARPCCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarpcqpagaalqadgrstctasaqscndlcehfcvnp 300
QY 301 DPGSTSCMCETGYRLAADOHRCEVDVDCILFSPPCPQRCVNTQSGFECHCYPNYDLVDG 360
DB 301 dpgstyscmcetgyrllaadghrcevdvdcilfepsppqcvtntqsgftechcypnydlvdg 360
QY 361 ECEVPDPCFRANCCEYQCPPLNOTSYLVCABGFAPLPHBPHRCQMFQNOTACPADCDPN 420
DB 361 ecevpdpfranceyqcpplntsylycvcaegfaplhpbhrcqmfqnotacpadcdpn 420
QY 421 TQASCECPGYIIDDGFICTDIDECENGFGSGVCHNLGTFPCTICGPPSALARHIGTDC 480
DB 421 tqascecpgyiiddgfictdidecengfgsgvchnlpgtfcticgppsalarhigtgc 480
QY 481 DSGKVDGSDSGSEPPSPTPGSTLTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgseppspptpgstltpavglvhs 516

RESULT 2
ID AAY83935
XX AAY83935 standard: Protein; 516 AA.

AC AAY83935;
XX
XX 28-JUL-2000 (first entry)
DE Human thrombomodulin TMD protein #2.
XX
XX Human; thrombomodulin; vasculitis; protein C; thrombin.
XX
XX Homo sapiens.
XX
XX JP2000053582-A.
XX
XX 22-FEB-2000.
XX
XX 06-AUG-1998; 98JP-0222688.
XX
XX 06-AUG-1998; 98JP-0222688.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX WPI; 2000-353249/31.
XX
XX N-PSDB; AAA10028.
XX
XX Treating agent for vasculitis contains peptide which promotes
XX activation of protein C by thrombin -
XX
PS Claim 4; Page 13-14; 18pp; Japanese.
XX
XX This sequence represents a human thrombomodulin protein. The invention
XX relates to a treating agent for vasculitis containing a peptide which
XX promotes activation of protein C by thrombin.
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVGALALAGLPAPAEPOPGSGQCEHDFALYPPGATFLNASQICDGLRGHLM 60
DB 1 mlgvlvgalalaglgfpapaepppgsgqcevdcfalypgpatflnaasqicdglrghlm 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPRRLGLRGFWVTGDNNTSYS 120
DB 61 tvrssvaadvysllngdgvgrrrlwigqlppgcgdprrlglrgfwvtdnntsys 120
QY 121 RMARLDLNGAPLPGPLCVANSAAEATVPSEPIWEBOQCEVKADGFLCEHFPATCRPLAY 180
DB 121 rmarldlngaplcgplcvansaeeatvpsepiweeqcevkadgflcehfpatcrplay 180
QY 181 EPGAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGIQLMCTAPPGAVGCHMARBP 240
DB 181 epgaaaavsitlygtpfaargadfgalpyvssaaavaplgqlmctappgavghmarbp 240
QY 241 GAWDCSVENGCEHACNAIPGARPCCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarpcqpagaalqadgrstctasaqscndlcehfcvnp 300
QY 301 DPGSTSCMCETGYRLAADOHRCEVDVDCILFSPPCPQRCVNTQSGFECHCYPNYDLVDG 360
DB 301 dpgstyscmcetgyrllaadghrcevdvdcilfepsppqcvtntqsgftechcypnydlvdg 360
QY 361 ECEVPDPCFRANCCEYQCPPLNOTSYLVCABGFAPLPHBPHRCQMFQNOTACPADCDPN 420
DB 361 ecevpdpfranceyqcpplntsylycvcaegfaplhpbhrcqmfqnotacpadcdpn 420
QY 421 TQASCECPGYIIDDGFICTDIDECENGFGSGVCHNLGTFPCTICGPPSALARHIGTDC 480
DB 421 tqascecpgyiiddgfictdidecengfgsgvchnlpgtfcticgppsalarhigtgc 480
QY 481 DSGKVDGSDSGSEPPSPTPGSTLTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgseppspptpgstltpavglvhs 516

Db 481 dsqkvdgdsqsgpppsptltpavglvhsq 516

RESULT 3

AA022016
ID AAR22016 standard; Protein; 516 AA.

AC AAR22016;

DT 03-JUL-1992 (first entry)

DE Truncated human thrombomodulin encoded by plasmid pSV2TMD1.

KW Thrombin binding site; blood clotting; TMD1 deleter.

OS Homo sapiens.

PN EP474273-A.

PD 11-MAR-1992.

PF 05-AUG-1991; 91EP-0202009.

PR 03-AUG-1990; 90JP-0204978.

PA (ASAH) ASAH KASEI KOGYO.

PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;

DR WPI; 1992-081820/11.

PT New polypeptide inhibits blood coagulation and platelet aggregation - promotes thrombin catalysed protein C activation for treating myocardial infarction, thrombosis, embolism, etc.

PS Example 1; Page 17; 112pp; English.

XX Plasmid M13mp19TMD1 (see AAR22013) was used as template for
CC site-directed mutagenesis. A 177bp fragment was deleted using the
CC "TMD1 deleter" oligonucleotide to give plasmid M13TMD1 which encodes
CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid
CC M13TMD1 was completely digested with HindIII and BamHI and a TMD1
CC fragment of ca. 1700bp was isolated. The fragment was ligated to
CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant
CC plasmid pSV2TMD1. See AAR22014-R22022 and AA025072.

XX Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 13; Length 516;
Best Local Similarity 99.8%; Pred. No. 7.5e-163;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVLYGALALAGLPAPAPAPGSGQCEVHDCFLALYGPATFLNASQICDGLRGLM 60
DB 1 mlyglvlgalaalaglpapapapgsgqcevhdcfalygpafllnasqicdglrghlm 60
OY 61 TVRBSVADYVSLINDGCGRRRLMIGLQPRCGDPRRLGRLGFWWTGNNNTSYS 120
DB 61 tvrsvadvysllndgvgvrrrlvlgqlppcgdprrlglrlgfwwtgndnts 120
OY 121 RWRALDNGAPLGLPLCAVAASAAEATVPSEPIWEQCEVAKDFLCEFFHPATCRPLAV 180
DB 121 rwarldngaplgplcvaasaaeatvpsepiweeqcevkadgflceffhpatcrplav 180
OY 181 EPGAAAANSTTGTTPPAARCAQFQALPVGSSAAVADLGLQMLCTAPGAVQGHAREAP 240
DB 181 epgaaaansttgttppaarcaqfualpvvgssaaavadlgqlmctapgavqghwareap 240
OY 241 GAWCSEVNGGCEHACNAIRGAPRCQCPAGALQADGRSCTASTGSCNLCHEFCVPPN 300
DB 241 gawcsevnggcehacnairgaprcqcpagalaqadgrsctastgscndlcehfcvppn 300

OY 301 DQPGSYSCMCEFGYRLAADOHRCEVDVDCILEPSPCQRCVNTQGFECCHCPNNTLYDGC 360
DB 301 dpgsyscmcefyrlaadohrcevdvdcillepsspqrvcvntqgfecchcpnyldvg 360
OY 361 ECEYEPDPCFRANCEYQCCPLNQTSTLYCAGGFAPIRHEPRRCQMFQNTACPADCPN 420
DB 361 eceyepdpcfranceyqccplnqstlycvaegfapihprhrcqmfncqtaepcdpn 420
OY 421 TQASCCEPEGYIIDDGFCITDIDECENGSGCVCNLTGFECTCGPDSALARIHTDC 480
DB 421 tqascecegyiiddgfcitdidecengsgfcsvchnlpgtfecicgpdosalvrhtdc 480
OY 481 DSGKVDGSDSGGSEPPSPPTSLTPPAVGLVHSG 516
DB 481 dsqkvdgdsqsgpppsptltpavglvhsq 516

RESULT 4

AA09347
ID AA09347 standard; Protein; 516 AA.

AC AA09347;

DT 08-JUL-1999 (first entry)

DE Human thrombomodulin SEQ ID NO:1.

KW Human; thrombomodulin; aqueous parenteral solution; storage;
KW distribution; acute coronary syndrome; thrombosis; embolism;
KW diabetes.

OS Homo sapiens.

PN WO918994-A1.

PD 22-APR-1999.

PF 13-OCT-1998; 98WO-JP04609.

PR 11-NOV-1997; 97JP-0308523.

PR 15-OCT-1997; 97JP-0281659.

PA (ASAH) ASAH KASEI KOGYO KK.

PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;

DR WPI; 1999-277444/23.

DR N-PSDB; AAX5879.

PT Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,
PT thrombosis, embolism, and diabetes

PS Claim 6; Page 87-89; 97pp; Japanese.

XX The present invention describes a method for maintaining the quality of
CC an aqueous parenteral solution of thrombomodulin comprising buffer and
CC surfactant aseptically filled in a case or syringe. Maintaining the
CC quality of an aqueous, parenteral thrombomodulin solution is
CC characterised in that the solution: (a) comprises soluble thrombomodulin,
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
CC filled into: (i) a case or (ii) a syringe without any empty space; and
CC (c) is kept in liquid form in storage and distribution and not frozen or
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome
CC (such as myocardial infarction, unstable angina and coronary artery
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
CC lupus erythematosus or Bartlett's syndrome), multiple organ failure,
CC disseminated intravascular coagulation, transient ischaemia, diabetes,
CC liver veno-occlusive diseases and deep vein thrombosis. The composition
CC is stable for a long period of time and can be stored and distributed in
CC ready to use form avoiding the problems of dissolution and accuracy when

CC preparing on demand. The present sequence represents human
CC thrombomodulin.
XX
SQ Sequence 516 AA;

50 Sequence 516 AA;

Query Match	99.9%	Score 2912;	DB 20;	Length 516;
Best Local Similarity	99.8%;	Pred. No. 7.5e-163;		
Matches 515; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	MATVIVIGALATAGGFEFAPAPAPQPGSSQCVENHCFFALYRGPRATFLNABSOICDILGRILHM	60
Dp	1	mlgvlvlgalaialag19fpapapqpgsqcvehcfcialygpactflnasqldcglrghlm	60
Qy	61	TVRSSVAADVDSLILNLDGSGVRRRLMIGLOIPGCDPDRGLRPLRGFOWTGMNNTSYS	120
Dp	61	tvrsvaadvdslilngdsgvrrrltwg19lppgcgdpkrlrglrgfwgdmntsys	120
Qy	121	RMARLDINCAPLICGPLCAVNSAAEATVVSSEPTWEEOCEVKADGFLCEBHFAPATCRPLAV	180
Dp	121	rwarldingaplicgplicvasaaetvvssep1weeqgcvekadgflcebfhfpatorplav	180
Qy	181	EPGAAAAAVSTVGFPEFARGADFALPVGSSAAVAPLGLQMLCSTAPPAVAGSHMARAP	240
Dp	181	epgaaaaavstvgfprfeargdafalpvgsaaavap19qlmctapspavagshmarap	240
Qy	241	GAMDCSYENSGCEHCNALIFGAPRCQCPAGALADGRSTASATQSCNDUCBHFVCPNP	300
Dp	241	gawdcsvenggcchacnal1pfgaprcqcpagaal1adgrstacsaatscndlcbhfcpnp	300
Qy	301	DOFQSYSCMERTVRIALAOHRCCEVDVDCILIEPSPCQQRVNNQGGFEOHCYAPNTDLVNG	360
Dp	301	dqfqsycmceertrylaadqhrccevdvdcil1eppcpqrvcvntggftechcypnvdlvng	360
Qy	361	ECVEPNDPCFRANCEYQCPPLNQTSTYLCVABEGFAP1PHERPHNCOMFNCOTACPADCPN	420
Dp	361	ecvepndpcfranceyqcpplnqstys1cvcabegfap1pberphncomfncntqacpadcpn	420
Qy	421	TQASCEPEEGYIINDDFICTDIDEEENGFCFSGVCHMLPETFECIGPSPALAAHRIITDC	480
Dp	421	tqascepegyi1nddfictdideengfcfsgvchml1pftfecigpsalavrhigtgc	480
Qy	481	DSGKVDGDSGSGEPPSPSTTPPAGVLVHSG	516
Dp	481	dsgrvdgdsgsgppspststltpavylvhsg	516

```

RESULT      5
ID          AAY83934
AC          AAY83934 standard; Protein; 516 AA.
XX
XX          AAY83934;
XX
DT          28-JUL-2000 (first entry)
XX
DE          Human thrombomodulin TMD protein.
XX
KW          Human; thrombomodulin; vasculitis; protein C; thrombin
XX
XX          Homo sapiens.
XX
OS          JP2000053582-A.
XX
PN          22-FEB-2000.
XX
PD          06-AUG-1998; 98JP-0222688.
XX
PF          06-AUG-1998; 98JP-0222688.
XX
PR          06-AUG-1998; 98JP-0222688.
XX
XX
PA          (ASAH ) ASAH KASEI KOGYO KK.
XX
XX
XX          WPI; 2000-353249/31.
XX          N-PSDB; AAA10027.
DR

```

XX Treating agent for vasculitis contains peptide which promotes
 PM activation of protein C by thrombin -
 PT
 XX
 PS Claim 4; Page 10-12; 18pp; Japanese.

CC This sequence represents a human thrombomodulin protein. The invention
CC relates to a treating agent for vasculitis containing a peptide which
CC promotes activation of protein C by thrombin.

SQ Sequence 516 AA;

Query Match	99.98%	Score 2912;	DB 21;	Length 516;
Best Local Similarity	99.88%	Pred. No. 7.5e-163;		
Matches 515; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	MGUVTCLAMLAGPFAPEAPQPGSQSCVEHHCOPALPEPAPFELNASOICOGIRCHIM	60
Db	1	mlgvltlglalaglgfipapeepqpgsqscvehhcfaipyapatfllnasqcdgltghlm	60
QY	61	TVRSSVADVLSLLINGSGGGRRLMLGLOLPGCGDPKRLGFLRGFOMVTDGDMNTSYS	120
Db	61	tvrsvvaadvlslllmgggygrtrlltvgldlppgcgdkprlglrgfqwtgdnnmsys	120
QY	121	RRARLDINGAPLCGFLCAVAASAEATVSRSEITMEQOCSEYKADSOFLCEHFRPATCPRLAY	180
Db	121	rwarldlngaprcgflcavasaeevpspseplweeqgevaxadglfclthpactprlav	180
QY	181	EPGAAAAVSVITYGTFPFARAKADFOALPVGSSAAVAPLGLJOLMCTAPGAVOCHMAREAP	240
Db	181	epgaaaaavsvltvgtpfarakadfoalpvgssaaavaprlgljlmctappgavghmareap	240
QY	241	GMDCSVENGCGEHCNMILPGAPROCCAGAAIADAGSCPSASATGSCNOLCEHFVCVBNP	300
Db	241	gmdcsvengcgchcnmlpgaprocagaaialdagscpsasatgscnolcehfvcvbnp	300
QY	301	DOPGYSQCMCEGTGYRLADONHRCEDVDVDCILPESPCPORCVNTGSGEHCNCTPNYDLVDG	360
Db	301	dpgysqcmcetgyrlaadqnhrcedvddcilpeppcpqrcvntggfchcnyplydlvdg	360
QY	361	ECVEEYVDFCFRANCHYQCPPLMNTSYLCLVCABEGAPRPHENHROMCNQTAGACDDPN	420
Db	361	ecveevpdcfranchyqcpplmntyylcvcaeagfarpbphrncqmtcnqtaacpddpn	420
QY	421	TQASCECEGGYTLIDGFCICTPIDECENGFGFSGVCHNLPTFFECICGPDLSALAHNITDC	480
Db	421	tqascecepgytlldgftictpidedeengfsgvchnlptgfecicgpdlsalvrhigtcd	480
QY	481	DSGKYVDGSDSGSEPPSPPTPESTITTPAVGLVHSG 516	
Db	481	dsghkvdygdsdsgeppspptpsttltppavylvhsg 516	

RESULT	6
AA69529	
ID	AA69529 standard; Protein; 516 AA.
XX	
AC	AA69529;
XX	
DT	10-APR-2000 (first entry)
XX	
DE	Human thrombomodulin variant, SEQ ID NO:3.
XX	
KW	Thrombomodulin; TM; soluble; affinity purification; cation exchange;
KW	anticoagulant; thrombolytic.
XX	
OS	Homo sapiens.
XX	
PN	JP11341990-A.
XX	
XX	
PD	14-DEC-1999.
XX	

PF 23-MAR-1999; 99JP-0077518.
 XX
 PR 30-MAR-1998; 98JP-0084389.
 XX
 PA (ASAHI) ASAH I KASEI KOGYO KK.
 XX
 WP: 2000-101696/09.
 DR N-PSDB; AA255965.
 XX
 PT Preparation of highly pure soluble thrombomodulin - used as an
 XX antibody coagulation agent and a thrombolytic agent
 XX
 PS Claim 9; Page 30-32; 36pp; Japanese.
 XX
 CC The invention relates to a novel method for the preparation of highly
 CC pure soluble thrombomodulin (TM) containing substantially no serum-
 CC derived or antibody-derived substance. The method comprises isolating
 CC soluble TM from unpurified serum via affinity purification using an
 CC anti-TM antibody. The soluble TM is then treated with with a cation
 CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to
 CC 4, and the fraction containing the soluble TM is isolated. The soluble
 CC TM can be used as a blood anticoagulation agent and a thrombolytic agent.
 CC This sequence represents a human thrombomodulin variant, designated
 XX SEQ ID NO:3.
 XX
 SO Sequence 516 AA;

QY 301 DQPSYSCMCEGTGYRLADDOHRCEDVDCLLEPSPCQRCVNTQGGFECHCYPNTDLYVG 360
 DB 301 dqpssycmcegtgyrlaadghrcedvdclleppspcqrcvntqggfdechcypnydlvdg 360
 QY 361 ECVPEVPDPCFRANCERYCQPLNOTSYLCVCAEGFAPLPHPHRCQMFQNTACRPADCPN 420
 DB 361 ecvpevpdpcfranceygcqplnqtsylcvcaegfaplhphrcqmfcncqtaacpdcn 420
 QY 421 TQASCEPPEGYLLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGSPSALARIHIGTDC 480
 DB 421 tqascepegyllddgfictdidecengfcsgvchnlpgtfecicgspalsalvrlhigtcd 480
 QY 481 DSGKVDGDSGSGEPSPPTPGSTLTTPRANGLVHSG 516
 DB 481 dsqkvdgdsqsgseppspptpgstlttpranglvhsg 516

RESULT 8
 AAR41806
 ID AAR41806 standard; peptide; 575 AA.
 AC AAR41806;
 XX 30-MAR-1994 (first entry)
 DT 30-MAR-1994 (first entry)
 DE Thrombomodulin.
 XX Transformation; fungus: blood coagulation; prevention; platelet;
 KM aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;
 KM site-directed mutagenesis; promotion; protein C; activation; thrombin.
 XX Homo sapiens.
 OS
 PN JP05213998-A.
 XX 24-AUG-1993.
 PD 24-AUG-1993.
 XX 02-AUG-1991; 91JP-0282369.
 PF 02-AUG-1991; 90JP-0204978.
 XX 03-AUG-1990; 90JP-0204978.
 PR 30-JUL-1991; 91JP-0189984.
 XX (ASAHI) ASAHI CHEM IND CO LTD.
 PA WPI: 1993-299652/38.
 XX Novel polypeptide obtd. by culturing transformed fungus - having
 PT blood coagulation preventing, platelet aggregation preventing and
 PR thrombolytic activities
 XX Disclosure; Fig 55; Japanese.
 PS
 XX Novel polypeptides, obtd. by culturing transformed fungus, have
 CC blood coagulation preventing, platelet aggregation preventing
 CC and thrombolytic activities.
 CC In an example, plasmid M13mp19MD3 (constructed from pSV2TM2
 CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected
 CC to site directed mutagenesis to prepare pSV2TM7. Plasmid pSV2TM7
 CC was transfected to COS-1 cells. The activity of promoting protein C
 CC activation by thrombin of the peptide produced by the transformed
 CC COS-1 cell was measured. The amt. of the peptide was determined.
 CC
 XX Sequence 575 AA;

Query Match 99.98; Score 2912; DB 14; Length 575;
 Best Local Similarity 99.88; Pred. No. 8.3e-163;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVVIVGALALAGLPAPAEPOPGSGQCVHDCFCALYPGATPLNASQICDGLRGHLM 60
 DB 1 mlgyvlgalalaglgfpapaepppgsgcvhdcfcaltypgatlfnasqicdglrgnlm 60

QY 61 TVRSSVAADVLSLLINGDGGVRRRLWGLQLPFGCGDPRFLRGLPQWMTGDNNTSYS 120
 DB 61 tvrsavaadvlsllingdggvrrrlwglqlpfgcgdprflrglpqlwmtgdnntsys 120
 QY 121 RWARLDLNGAPLCGFLCYAASAAEATVPSPTWEEOCEKADGFLCERHFPATCRPLAV 180
 DB 121 rwarldlngaplcgflcyavaaaatvpsptweeocekadgflcerhfpatcrplav 180
 QY 181 EFGAAAAAIVTYGTPFARAGADFOALPVGSSAAVAFGLQMLMCTAPPGAVOGHMAAREAP 240
 DB 181 efgaaaaaivtygtpfaragadfoalpvgssaaavafglqmlmctappgavoghmareap 240
 QY 241 GAMDCSVENGGEHACNMLIPGARPCQCPAGALDADRSTASTATQSCNDLCEHCEVVRP 300
 DB 241 gamdcsvengehacnmlipgarpcqcpagaldadrstastatqscndlcehcevvrrp 300
 QY 301 DQPSYSCMCEGTGYRLADDOHRCEDVDCLLEPSPCQRCVNTQGGFECHCYPNTDLYVG 360
 DB 301 dqpssycmcegtgyrlaadghrcedvdclleppspcqrcvntqggfdechcypnydlvdg 360
 QY 361 ECVPEVPDPCFRANCERYCQPLNOTSYLCVCAEGFAPLPHPHRCQMFQNTACRPADCPN 420
 DB 361 ecvpevpdpcfranceygcqplnqtsylcvcaegfaplhphrcqmfcncqtaacpdcn 420
 QY 421 TQASCEPPEGYLLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGSPSALARIHIGTDC 480
 DB 421 tqascepegyllddgfictdidecengfcsgvchnlpgtfecicgspalsalvrlhigtcd 480
 QY 481 DSGKVDGDSGSGEPSPPTPGSTLTTPRANGLVHSG 516
 DB 481 dsqkvdgdsqsgseppspptpgstlttpranglvhsg 516

RESULT 9
 AAR43031
 ID AAR43031 standard; protein; 575 AA.
 AC AAR43031;
 XX 16-MAY-1994 (first entry)
 DT 16-MAY-1994 (first entry)
 DE Human thrombomodulin.
 XX Anticoagulant; platelet aggregation inhibitor;
 KM protein C; activation; thrombin; thrombomodulin;
 KM coagulation disorder; thrombolysis; myocardial infarction;
 KM embolism; telangiectasis; arteriosclerosis obliterans;
 KM disseminated intravascular coagulation; DIC; angina pectoris;
 KM gestosis; transient ischaemic attack.
 XX
 OS Homo sapiens.
 PN WO9322447-A.
 XX 11-NOV-1993.
 PD 11-NOV-1993.
 XX 30-APR-1993; 93WO-JP00578.
 PF 30-APR-1993; 92JP-0112903.
 XX 01-MAY-1992; 92JP-0112903.
 PR (ASAHI) ASAHI CHEM IND CO LTD.
 PA (ASAHI) ASAHI KASEI KOGYO KK.
 XX Kondo S, Toma K, Zushi M;
 DR WPI: 1993-368806/46.
 XX Peptide with anticoagulant and platelet aggregation inhibitor
 PT activity - which promotes protein C activation by thrombin and is
 PT useful in treating coagulation disorders e.g. thrombosis
 XX Disclosure; Fig 1; 84pp; Japanese.

XX New peptides (see AAR50069) are inhibitors of the blood coagulation
 CC and platelet aggregation activities of thrombin and promote the
 CC protein-C activation effect of thrombin. They can be produced
 CC efficiently in pure form by culture of appropriate transformants,
 CC and are useful in treatment of circulatory disorders such as
 CC myocardial infarction, thrombosis, embolism, telangiectasis,
 CC arteriosclerosis obliterans, disseminated intravascular
 CC coagulation, angina pectoris, gestosis and transient ischaemic
 CC attack.

Sequence 575 AA:

Query Match 99.9%; Score 2912; DB 14; Length 575;
 Best Local Similarity 99.8%; Pred. No. 8.3e-163;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVIVLGLALAGLGFPAAPAPPGSGSCVEHDCFALYPGATFLNASQICDGLRGLM 60
 DB 1 mlgvlylgalaialaglfpaapepgsgscvehdcfalyppatflnasqicdglrghlm 60
 OY 61 TVRSSVADYISLLNGDGVGRRRLMIGLQLPGCGDKPKRLGFGQWVTGDNNTSYS 120
 DB 61 tvrssvaadvysllngdgvgrrrlmigqlpgcgdkprkrlgfgwvtgdnntsys 120
 OY 121 RMARLDLNGAPLGPICLVANSAEATVPSEPIWEEOGCEVADGFLCEHFHPCRPRLAY 180
 DB 121 rmarldlngaplgpiclvansaeatvpsepiweeogcevadgflcefhfpcrprlav 180
 OY 181 EPGAAAAVSTYGTTPFAAGADFOALPVSSAAVAPLGLQMLCTAPPAVGHMARBP 240
 DB 181 epgaaaaavstlygttppfaagadfoalpvssaaavaplgqlmctappavghmarbp 240
 OY 241 GAWDCSVENGCEHACNAIPGAPRCOCAPGALQADGSCGTAASQSCNDICEHFCVNP 300
 DB 241 gawdcsvengcehacnaipgaprcocapgalqadgscgtaasqscndicehfcvnp 300
 OY 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILBSPPCORCVNTOGGFECHCYPNYDLVG 360
 DB 301 dpgsyscmcegyrlaadqhrcevdvdcilbsspccorcvntoggfechcypnydlvg 360
 OY 361 ECVPEVDFCFRANCEYQCQPLNOTSYLCAEGFAPRIEHRRCOMFCNOTACPADCDPN 420
 DB 361 ecvpevdpcfanceyqcqplnotsylycægfpapriehrcomfcnotacpadcdpn 420
 OY 421 TQASCECPGTYLDDGFTDIDECENGFCGVCNHLPGTEFCICGPDASALARIHIGDC 480
 DB 421 tqascecpgyllddgftdiddecengfcgvcnhlpgtefcicgpdasalarihigdc 480
 OY 481 DSGKVDGSDSGSEPPSPSTLTPPAVGLVHSG 516
 DB 481 dsqkvdgdsdsgseppspstltpavglvhsq 516

RESULT 10

AAR10617 standard; Protein; 515 AA.

AC AAR10617;

XX 17-APR-1991 (first entry)

DE Soluble thrombomodulin deriv.

XX Thrombosis; anticoagulant; Protein C.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..16 /label= signal peptide

FT /note= "For low mol. wt. deriv.; optionally present"

FT Peptide

FT 1..18

/label= signal peptide

/note= "for high mol wt. deriv.; optionally present"

XX EP412841-A.

XX 13-FEB-1991.

XX 10-AUG-1990; 90EP-0308626.

XX 05-FEB-1990; 90US-0474870.

XX 11-AUG-1989; 89US-0393617.

XX (ELI) ELI LILLY & CO.

XX Bang NU, Grinnell BW, Hoskins JA, Moore RE, Parkinson JF;

XX WPI; 1991-045960/07.

XX N-PSDB; AAQ10435.

PT New soluble thrombomodulin derivatives - produced by recombinant
 PT DNA techniques for use as anticoagulants and in treatment of
 PT thrombosis.

XX Claim 2; Page 45; 81pp; English.

The sequence is that of a sol. thrombomodulin deriv. comprising the
 CC signal peptide (optional), N-terminal, epidermal growth factor
 CC homology region and the Ser/Thr-rich region, but lacking the trans-
 CC membrane and cytoplasmic domains. The sequence was deduced DNA
 CC obt'd. from clone GH73A isolated from a human chromosome 20 library,
 CC subcloned into pUC19 to give pGH73A. A Pvu MI fragment of this
 CC vector was ligated with a linker to construct pUC187M, which was
 CC treated with BamI to delete about 500 bp from the 3' end of the
 CC gene. (The deletion occurs at the Ser/Thr-rich/transmembrane domain
 CC junction, effectively cleaving the transmembrane and cytoplasmic
 CC domains.) The plasmid was recircularised to give pUC187MD for use
 CC in prodn of the recombinant sol. deriv. The protein reacts with
 CC thrombin to activate the Protein C anticoagulant pathway, inhibit
 CC thrombin, and can be used to treat or prevent thrombotic disorders.

XX Sequence 515 AA:

Query Match 99.8%; Score 2910; DB 12; Length 515;
 Best Local Similarity 100.0%; Pred. No. 9.8e-163;
 Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGVIVLGLALAGLGFPAAPAPPGSGSCVEHDCFALYPGATFLNASQICDGLRGLM 60
 DB 1 mlgvlylgalaialaglfpaapepgsgscvehdcfalyppatflnasqicdglrghlm 60
 OY 61 TVRSSVADYISLLNGDGVGRRRLMIGLQLPGCGDKPKRLGFGQWVTGDNNTSYS 120
 DB 61 tvrssvaadvysllngdgvgrrrlmigqlpgcgdkprkrlgfgwvtgdnntsys 120
 OY 121 RMARLDLNGAPLGPICLVANSAEATVPSEPIWEEOGCEVADGFLCEHFHPCRPRLAY 180
 DB 121 rmarldlngaplgpiclvansaeatvpsepiweeogcevadgflcefhfpcrprlav 180
 OY 181 EPGAAAAVSTYGTTPFAAGADFOALPVSSAAVAPLGLQMLCTAPPAVGHMARBP 240
 DB 181 epgaaaaavstlygttppfaagadfoalpvssaaavaplgqlmctappavghmarbp 240
 OY 241 GAWDCSVENGCEHACNAIPGAPRCOCAPGALQADGSCGTAASQSCNDICEHFCVNP 300
 DB 241 gawdcsvengcehacnaipgaprcocapgalqadgscgtaasqscndicehfcvnp 300
 OY 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILBSPPCORCVNTOGGFECHCYPNYDLVG 360
 DB 301 dpgsyscmcegyrlaadqhrcevdvdcilbsspccorcvntoggfechcypnydlvg 360
 OY 361 ECVPEVDFCFRANCEYQCQPLNOTSYLCAEGFAPRIEHRRCOMFCNOTACPADCDPN 420

Db 361 ecvepvpdcfranceygcqplnqtsylcvcaeagfapipheprcmfcncqtaacpdcn 420
 QY 421 TOASCCEPEGYILDGFCITDIDECENGCGSCVCHNLDEFTFCICGPPSALARIHGTDC 480
 Db 421 tqascecepegyildgfcitdidecengcgscvchnlpgflectcpgdsalarhgtac 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPPAVGLVHS 515
 Db 481 dsqkvdgdsqsgseppspptgsltpavglvhs 515

RESULT 11

AA82070
 ID AAP82070 standard; protein; 575 AA.

AC AAP82070;

DT 19-OCT-1990 (first entry)

DE Human thrombomodulin encoded by plasmid p2.1.

KW thrombomodulin activity; protein C; anticoagulant;

RW epidermal growth factor (EGF) domains.

OS sythetic.

PN MO8809811-A.

PD 15-DEC-1988.

PE 09-JUN-1988; 88WO-DK00089.

PR 12-JUN-1987; 87DK-0002990.

PA (NOVO) NOVO INDUSTRI A/S.

PI Nexo BA, Esper B;

DR WPI: 1988-36826/51.

DR N-PSDB; AAP82026.

PT Recombinant protein having thrombomodulin activity -

PT used in the therapeutic control of coagulation and the treatment

PT and prevent of thrombotic episodes

PS Disclosure; ; P; English.

CC Protein is encoded by plasmid p2.1 derived from human cell

CC line A549 known to express about 10000 molecules of thrombomodulin

CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a

CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to

CC potentiate a patient's anticoagulant capacity.

CC See also AAN82027 and AAN82037.

CC Sequence 575 AA;

Query Match 99.8%; Score 2910; DB 9; Length 575;

Best Local Similarity 99.8%; Pred. No. 1.1e-162;

Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFAPAEPOPGSGQVEHDCFAIYPGATPLNAGQIDGRLGHLM 60
 Db 1 mlgvlvlgalaaglgfapaeppgsgqvehdcafyppgatlfnasqicdglrghlm 60
 QY 61 TVRSSVADVISLLNDSGVGRRRLWIGLQLEPGCGDKRUGLPGFQWVGDNNTSYS 120
 Db 61 tvrssvadvisllndsgvgrrrlwiglqlpepgcgdkrulgplgfwvgtgdnntsys 120
 QY 121 RWAFLDINGAPLIGPLVAVSAEAATPSPITWEEQCCVADGRTCEHFRATGRPLAV 180
 Db 121 rwafldingapligplvavsaatpsepitweeqcevkadaficethfratcprplav 180

QY 181 EPGAAAAVSYTYGTFEFAARGADFOALPVSSAAVAPLGLIMCTAPGAVOGHAREAP 240
 Db 181 epgaaaaavsytygtfeffaargadfoalpvssaaavaplgilmtctappgavoghmareap 240
 QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGSCSTASXTQSCNDLCERFCYVNP 300
 Db 241 gamdcevngecehacnalpgaprcqcpagallqadgscstasxtqscndlcerfcyvpnp 300
 QY 301 DQGSYSCMCEGYRLAADORCEBDVDCITLSPSCPPRCVNTQAGCEHCCHPYVDLVDG 360
 Db 301 dpgysyrcmcegyrlaadorcbedvdcitlspscpprcvntqagcehcchpyvdlvdg 360
 QY 361 ECVEPVDPCEFRANCEYOCPLNQTSTYLCVCAEGFAPIPHEPRCMFCNCQTAACPDN 420
 Db 361 ecvepvpdcfranceygcqplnqtsylcvcaeagfapipheprcmfcncqtaacpdcn 420
 QY 421 TOASCCEPEGYILDGFCITDIDECENGCGSCVCHNLDEFTFCICGPPSALARIHGTDC 480
 Db 421 tqascecepegyildgfcitdidecengcgscvchnlpgflectcpgdsalarhgtac 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPPAVGLVHS 516
 Db 481 dsqkvdgdsqsgseppspptgsltpavglvhs 516

RESULT 12

AA82018
 ID AAR22018 standard; Protein; 516 AA.

AC AAR22018;

DT 03-JUL-1992 (first entry)

DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.

KW Mutant; thrombin binding site; blood clotting; Tm2 mutator.

OS Homo sapiens.

PN EP474273-A.

PD 11-MAR-1992.

PE 05-AUG-1991; 91EP-0202009.

PR 03-AUG-1990; 90JP-0204978.

PA (ASAHI) ASahi KASEI Kogyo.

PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;

DR WPI: 1992-081820/11.

PT New polypeptide inhibits blood coagulation and platelet

PT aggregation - promotes thrombin catalysed protein C activation

PT for treating myocardial infarction, thrombosis, embolism, etc.

PS Example 1; Page 18; 112pp; English.

CC Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal
 CC amino acids of human thrombomodulin was used as a template for
 CC site-directed mutagenesis using the "Tm2 mutator" to produce
 CC plasmid M13TMM2. In the mutant thrombomodulin encoded by the
 CC plasmid, the wild-type Asp residue at position 367 is substituted
 CC by an Glu residue. The activity of this mutant, truncated
 CC thrombomodulin (i.e. D123Glu) in the activation of protein C was
 CC compared to that of similarly truncated thrombomodulin with Asp
 CC at position 367 (i.e. D123Asp). The activity of D123Glu was as high
 CC as 1.2 times that of D123Asp. See AAR22013-R22022 and AAQ25074.

Sequence 516 AA;

Query Match 99.7%; Score 2908; DB 13; Length 516;
Best Local Similarity 99.6%; Pred. No. 1.3e-162;
Matches 516; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MGVLVLAGALAGLPAPAEPOPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlgvlylgalaiaaglfapaeppgsgqvehdcfalypgpattflnasqicdglrghlm 60
QY 61 TVRSSVADVISTLLNGGCGVRRRLMIGLQLPCCGPKRLGRLRGQWVTGDNNTSYS 120
DB 61 tvrsadvadvistllnggvgvrrrlwlgqlppcgdkrlgprlgfvgvtdnntsys 120
QY 121 RMARLDLNGAPLCGPLCVASAATVPSEPIWEQCEVADRGLCFEHPFATCRPLAY 180
DB 121 tmarldlngaplcgplcvasaaatvpsepiweeqcevakdglficehfpatcrplay 180
QY 181 EPGAANAASITVGPFAARGADFOALPVSSAAVAPLGLQLMCTAPPGAVGHWAREAP 240
DB 181 epgaanaasitvgtpfaragadfgalpyvssaaavaplgqlmctappgavghwareap 240
QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvenggcchacnaipgarcqcpagalaqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDDCILPSPPCORCVNTQGFECICYPNYDLVDG 360
DB 301 dqpgyscmcegyrllaadqhrcevdcdcllpspcqrcvntqgfyecicypnydlvdg 360
QY 361 ECVEPVDFCFRANCEYOCOPLNQTSYLCAEGFAPRPHRRCOMFCNOTCPADCDPN 420
DB 361 ecvepvdfcfranceyocqplnqtsylcvcaeafaprhprcqmfcncqtacpadcdpn 420
QY 421 TQASCECPGYIILDGFICTDIDECENGFCGVCVCHNLPGTEFCICGPDALARIHIGTDC 480
DB 421 tqascecpgyiildgfiictdidceengfcsgvchnlpgtfeccicgpdalarihigtcd 480
QY 481 DSGKVDGSDSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgpeppspptpgstltppavglvhs 516
```

RESULT 13

AAR31572 standard; Protein: 575 AA.

AAR31572;

27-MAY-1993 (first entry)

Human thrombomodulin.

Mutagenesis: site specific; nucleic acid constructs;
restriction site; introduction; removal.

Synthetic.

Location/Qualifiers

/note= "signal peptide"

/note= "EGF-1"

/note= "EGF-2"

/note= "EGF-3"

/note= "EGF-4"

/note= "EGF-5"

/note= "EGF-6"

FT /note= "O-linked glycosylation domain"
FT 516..538
FT /note= "stop transfer sequence"
FT 539..575
FT Domain
FT /note= "cytoplasmic domain"

MO9301282-A.
21-JAN-1993.
01-JUL-1992; 92WO-0505573.
PR 01-JUL-1991; 91US-0724237.
PA (BERL-) BERLEX LAB INC.
PI Andrews WH, Morser MJ, Villander LR;
DR WPI: 1993-045488/05.
XX Site-specific mutagenesis of nucleic acid constructs - using an
XX oligo:nucleotide which changes a nucleotide and introduces or
XX removes a restriction site
XX Example: fig 3; 87pp; English.
CC The sequence is that of the native human thrombomodulin showing the
CC six EGF-like domains.
XX
SQ Sequence 575 AA.

Query Match 99.7%; Score 2908; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.4e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MGVLVLAGALAGLPAPAEPOPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlgvlylgalaiaaglfapaeppgsgqvehdcfalypgpattflnasqicdglrghlm 60
QY 61 TVRSSVADVISTLLNGDGVGRRRLMIGLQLPCCGDKRLGRLFGQWVTGDNNTSYS 120
DB 61 tvrsadvadvistllngdgvvrrrlwlgqlppcgdkrlgprlgfvgvtdnntsys 120
QY 121 RMARLDLNGAPLCGPLCVASAATVPSEPIWEQCEVADRGLCFEHPFATCRPLAY 180
DB 121 tmarldlngaplcgplcvasaaatvpsepiweeqcevakdglficehfpatcrplay 180
QY 181 EPGAANAASITVGPFAARGADFOALPVSSAAVAPLGLQLMCTAPPGAVGHWAREAP 240
DB 181 epgaanaasitvgtpfaragadfgalpyvssaaavaplgqlmctappgavghwareap 240
QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvenggcchacnaipgarcqcpagalaqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDDCILPSPPCORCVNTQGFECICYPNYDLVDG 360
DB 301 dqpgyscmcegyrllaadqhrcevdcdcllpspcqrcvntqgfyecicypnydlvdg 360
QY 361 ECVEPVDFCFRANCEYOCOPLNQTSYLCAEGFAPRPHRRCOMFCNOTCPADCDPN 420
DB 361 ecvepvdfcfranceyocqplnqtsylcvcaeafaprhprcqmfcncqtacpadcdpn 420
QY 421 TQASCECPGYIILDGFICTDIDECENGFCGVCVCHNLPGTEFCICGPDALARIHIGTDC 480
DB 421 tqascecpgyiildgfiictdidceengfcsgvchnlpgtfeccicgpdalarihigtcd 480
QY 481 DSGKVDGSDSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgpeppspptpgstltppavglvhs 516
```

```

RESULT 14
ID AAR22017 standard; Protein; 516 AA.
XX
AC AAR22017;
XX
DT 03-JUL-1992 (first entry)
XX
DE Human thrombomodulin (1-516) with Asp367 substituted by Ala.
XX
KM Mutant; thrombin binding site; blood clotting; Tmml mutator.
XX
OS Homo sapiens.
XX
PN EP474273-A.
XX
PD 11-MAR-1992.
XX
PF 05-AUG-1991; 91EP-0202009.
XX
PR 03-AUG-1990; 90JP-0204978.
XX
PA (ASAH ) ASAH KASEI KOGYO.
XX
PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
XX
DR WPI; 1992-081820/11.
XX
PT New polypeptide inhibits blood coagulation and platelet
PT aggregation - promotes thrombin catalysed protein C activation
PT for treating myocardial infarction, thrombosis, embolism, etc.
XX
PS Example 1; Page 18; 112pp; English.
XX
CC Plasimid M13TMD1 (see AAR22016) encoding the first 516 N-terminal
CC amino acids of human thrombomodulin was used as a template for
CC site-directed mutagenesis using the "Tmml mutator" to produce
CC Plasimid M13TMD1. In the mutant thrombomodulin encoded by the
CC plasimid, the wild-type Asp residue at position 367 is substituted
CC by an Ala residue. This mutant, truncated thrombomodulin has low
CC activity in the activation of protein C, c.f. high activity for the
CC similarly truncated thrombomodulin with Asp 367.
CC See AAR22013-R22022 and AAQ25073.
XX
SQ Sequence 516 AA;

Query Match 99.6%; Score 2904; DB 13; Length 516;
Best Local Similarity 99.6%; Pred. No. 2.2e-162;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MGVLVLGALALAGLGFPAAPBPQPGSGVCEHDCFALYPGATFTINAQICDGLGHLM 60
DB 1 MGVLVLGALALAGLGFPAAPBPQPGSGVCEHDCFALYPGATFTINAQICDGLGHLM 60
OY 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPPGCGDKRRLGLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPPGCGDKRRLGLRGFQWVTGDNNTSYS 120
OY 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPPGCGDKRRLGLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPPGCGDKRRLGLRGFQWVTGDNNTSYS 120
OY 121 RKAARDLNCAPLCGLCVAAEAATVPSEPIWEEOCEKADGFLCEFHFPATCRPLAY 180
DB 121 RKAARDLNCAPLCGLCVAAEAATVPSEPIWEEOCEKADGFLCEFHFPATCRPLAY 180
OY 121 RKAARDLNCAPLCGLCVAAEAATVPSEPIWEEOCEKADGFLCEFHFPATCRPLAY 180
DB 121 RKAARDLNCAPLCGLCVAAEAATVPSEPIWEEOCEKADGFLCEFHFPATCRPLAY 180
OY 181 EFGAAAAANVTITGTFPAARAGADFOALPYGSSAANAPLGLQIMCTARPGAVGHWAREAP 240
DB 181 EFGAAAAANVTITGTFPAARAGADFOALPYGSSAANAPLGLQIMCTARPGAVGHWAREAP 240
OY 241 GAMDCSVEGGEHCACNAIPGAPRCQPGALQADGRSCTASANTOSCDLCHFCVPP 300
DB 241 GAMDCSVEGGEHCACNAIPGAPRCQPGALQADGRSCTASANTOSCDLCHFCVPP 300
OY 301 DPGSYSCMCTGYRLAADQHRCEVDQCTLEPSPCQRCVMTQSGFECGHCYNYPLVNG 360
DB 301 DPGSYSCMCTGYRLAADQHRCEVDQCTLEPSPCQRCVMTQSGFECGHCYNYPLVNG 360

```

```

DB 301 dpgsyscmctgyrllaadqhrcevdqctleppspcpqrcvmtqsgftechcypnydlvdg 360
OY 361 ECVPEVDPFCFRANCEYCCQLNGTSTYLCVCAEGFAPLPHEPHRCOMFCNQTACPADCDPN 420
DB 361 ECVPEVDPFCFRANCEYCCQLNGTSTYLCVCAEGFAPLPHEPHRCOMFCNQTACPADCDPN 420
OY 421 TOASCECPGEGYIIDDGFICNDIDCECGSGVCNHLPTGPFICGPDGALARRHGTDC 480
DB 421 TOASCECPGEGYIIDDGFICNDIDCECGSGVCNHLPTGPFICGPDGALARRHGTDC 480
OY 481 DSGKVDGDSGSGEPSPPTGSTLTPPAVGLVHSG 516
DB 481 dsgkvdgdsdsgsgppspptgstltppavglvhs 516

RESULT 15
ID AAR20639 standard; Protein; 575 AA.
XX
AC AAR20639;
XX
DT 01-MAY-1992 (first entry)
XX
DE Human urinary thrombomodulin.
XX
KM Blood clotting; intravascular coagulation.
XX
OS Homo sapiens.
XX
PS Key Location/Qualifiers
PS Peptide 1..18
PS FT /label= signal
PS FT 19..575
PS FT /label= thrombomodulin
XX
PN WO9200325-A.
XX
PD 09-JAN-1992.
XX
PF 27-JUN-1991; 91WO-JP00873.
XX
PR 27-JUN-1990; 90JP-0168766.
XX
PA (MOCH ) MOCHIDA PHARM KK.
XX
PI Nii A, Morishita H, Uemura A, Mochida E;
XX
DR WPI; 1992-041517/05.
DR N-PSDB; AAQ20810.
XX
PT New modified recombinant human urinary thrombomodulin - has
PT thrombin binding, anticoagulant and thrombolytic activity and is
PT for treating blood coagulation disorders
XX
PS Claim 1; Fig 3; 107pp; Japanese.
XX
CC A 2.5kb cDNA fragment was isolated by screening a human placental
CC cell library with a probe coding for the N-terminal of human
CC thrombomodulin. This full-length sequence was modified to produce
CC two alternative, truncated forms of the coding sequence encoding
CC only up to amino acid 456 of the mature polypeptide: the Ala
CC residue at position 455 is substituted by Val in one of the two
CC truncated forms.
XX
SQ Sequence 575 AA;

Query Match 99.6%; Score 2903; DB 13; Length 575;
Best Local Similarity 99.6%; Pred. No. 2.8e-162;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MGVLVLGALALAGLGFPAAPBPQPGSGVCEHDCFALYPGATFTINAQICDGLGHLM 60
DB 1 MGVLVLGALALAGLGFPAAPBPQPGSGVCEHDCFALYPGATFTINAQICDGLGHLM 60

```

Db 1 mlgvlylgatlaglgfpapaeppgsgqcvenhcfalypgpafliinasqldqlrgqlm 60
QY 61 TVSSSVADYISLLNGDGVGRRRLMIGQLPFGCGDPRKRLGPFQWNTGDNNTSYS 120
Db 61 tvsssvadvysllngdvgvgrrrlwlglqlppgcgdkrlgplrgfwltdgnntsys 120
QY 121 RMARLDINGAPLPGPLCVANSAATVPSEPTWEEQCEVKADGFLCEHFPATCRPLAV 180
Db 121 rwarldingaplcgplcvavsaatvpseptweeqcevkadgflcefhfpatcrplav 180
QY 181 EPGAAAAVITGTTPAARGADROALPVGSSAAVAPLGIQLMCTARPGAVQGHAREAP 240
Db 181 epgaaaaavsltygtptaargadlqalpvgsaaavaplgqlmctappgavqghareap 240
QY 241 GANDCSYENGSCHEHACNAIPGAPRCOPAGALQADGRSCTASATOSCNDLCEHPCVNP 300
Db 241 gawdcsvengscelhacnaipgaprcqpagaalqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSTSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTOGFECHCYPNYDLVDG 360
Db 301 dqpgystcmctgyrlaadqhrcevdvdcillepspcrcvntogfchcypnydlvdg 360
QY 361 ECEVPDPCFRANCEYOCOPLNOTSTYLCVCAEGFAPTPHEPHRCQMFCNOTACPADCDPN 420
Db 361 ecevpdpctranceyocqplnqstylcvcaegfaptphephrcqmfcnqtaacadpn 420
QY 421 TQASCECEPEGYILDDGFICTDIDECENGSGVCHNLPGTFECICGPDLSALARIHTDC 480
Db 421 tqascecepegyliddgfictdidecengsfcsychnlpgtfecicgpdlsalarhltcdc 480
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516
Db 481 dsqkvdgdsqsggeppspptpgstltppavglvhsq 516

Search completed: May 8, 2002, 12:38:54
Job time: 215 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:54 ; Search time 29.56 Seconds
(without alignments)
1329.705 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGVLVGLALAGLGFPPAP.....PSPPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR:68:***
2: PIR:1:***
3: PIR:2:***
4: PIR:4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	575	1 THH01	thrombomodulin pre
2	1899	65.1	577	2 A60501	thrombomodulin pre
3	1092.5	37.5	356	2 A25918	thrombomodulin - b
4	416	14.3	1620	2 T27283	hypothetical prote
5	414	14.2	1574	2 T13554	MEGF6 protein - ra
6	373	12.8	2907	2 A57278	fibriillin-2 precu
7	367.5	12.6	1184	2 A55184	fibriillin-2 precu
8	367	12.6	2871	2 A55184	fibriillin-2 precu
9	365	12.5	2918	2 A54105	fibriillin-2 precu
10	363	12.4	3002	2 A47221	fibriillin-1 precu
11	361	12.4	2871	2 A55624	fibriillin-1 precu
12	359	12.3	741	2 T46488	hypothetical prote
13	357	12.2	1221	2 A49457	fibulin-2 precu
14	331	11.4	589	2 T43210	fibulin-1D precu
15	329.5	11.3	1964	2 T09059	notch4 - mouse
16	327.5	11.2	1712	2 A38261	masking protein pr
17	324.5	11.1	689	2 T42760	fibulin, splice fo
18	324.5	11.1	712	2 T42990	fibulin 1, splice
19	320.5	11.0	2321	2 S78586	notch3 protein - h
20	320	11.0	1394	2 A35636	transforming grow
21	314.5	10.8	3507	2 T34513	hypothetical prote
22	313	10.7	601	2 B36346	fibulin 1 precu
23	313	10.7	683	2 C36346	fibulin 1 precu
24	313	10.7	798	2 T22793	hypothetical prote
25	312.5	10.7	1820	2 A55494	latent transformin
26	308	10.6	685	2 S78040	fibulin, splice fo
27	308	10.6	705	2 S74968	fibulin, splice fo
28	303.5	10.4	1251	2 A57293	latent transformin
				Notch-1 protein -	

30	286	9.8	2471	2 A49128	cell-fate determin
31	285	9.8	2352	2 T30201	Notch homolog prot
32	283.5	9.7	2437	2 S42612	transmembrane prot
33	280.5	9.6	2703	1 A24420	notch protein - fr
34	280	9.6	2318	2 S45306	notch 3 protein -
35	278.5	9.5	387	2 I38449	extracellular prote
36	277.5	9.5	511	2 T17298	hypothetical prote
37	277	9.5	2531	2 S18188	notch protein homo
38	276.5	9.5	493	2 JC5621	epidermal growth f
39	272	9.3	2555	2 A40043	notch protein homo
40	269.5	9.2	1217	1 EGMSNG	epidermal growth f
41	263	9.0	1203	2 A49175	Notch B protein -
42	259.5	8.9	2531	2 T31070	notch homolog - se
43	256.5	8.8	2524	2 A35844	notch protein - Af
44	254.5	8.7	674	2 I55476	growth potentialin
45	249.5	8.6	1133	1 EGRT	epidermal growth f

ALIGNMENTS

RESULT 1
THH01
thrombomodulin precursor (validated) - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 15-Sep-2000
C/Accession: A41442; A28307; A29680; A27073; JX0264; 538954
R/Shiomi, T.; Shiomi, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Mar
J. Biochem. 103, 281-285, 1988
A>Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed
A:Reference number: A41442; MUID:88227901
A:Accession: A41442
A:Molecule type: DNA
A:Residues: 1-575 <SH1>
A:Cross-references: DBJ:J02010; NID:9220126; PIDN:BA00149.1; PID:9220127
R/Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987
A>Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of t
A:Reference number: A28307; MUID:87317665
A:Accession: A28307
A:Molecule type: DNA
A:Residues: 1-472, 'A', 474-575 <JNC>
A:Cross-references: GB:J02973; NID:9339656; PIDN:AAA61175.1; PID:9339659
R/Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioke, J.; Maruyama, T.; Kushi, M.;
EMBO J. 6, 1891-1897, 1987
A>Title: Structure and expression of human thrombomodulin, a thrombin receptor on
A:Reference number: A29680; MUID:88004395
A:Accession: A29680
A:Molecule type: mRNA
A:Residues: 1-575 <SU2>
A:Cross-references: GB:X05495; NID:937123; PIDN:CA29045.1; PID:9736251
A:Experimental source: Lung endothelium
A:Note: part of this sequence, including the amino end of the mature protein, were
R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.
Biochemistry 26, 4350-4357, 1987
A>Title: Human thrombomodulin: complete cDNA sequence and chromosome localization
A:Reference number: A27073; MUID:88024950
A:Accession: A27073
A:Molecule type: mRNA
A:Residues: 1-472, 'A', 474-575 <WEN>
A:Cross-references: GB:M16552; NID:9339656; PIDN:AMB59508.1; PID:9339657
A:Experimental source: Placenta
A:Note: parts of this sequence were determined by protein sequencing
R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Okuchi, M.; Kimura, S.; Aoki, N.
J. Biochem. 113, 433-440, 1993
A>Title: Urinary thrombomodulin, its isolation and characterization.
A:Reference number: JX0264; MUID:93293792
A:Accession: JX0264
A:Molecule type: protein; mRNA
A:Residues: 19-472, 'A', 474-486 <YAM>
A:Experimental source: urine
A:Note: the urinary form appears to be identical with that circulating in plasma
R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell,

Biochem. J. 295, 131-140, 1993
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble
serine.
A:Reference number: S38954; MUID:94029900
A:Accession: S38954
A:Molecule type: protein
A:Residues: 475-491/X; 493-494 <GER>
A:Note: The residue designated 'X' was determined to be a Ser with covalently bound choro
R;Malminger, D.P.; Komives, E.A.
A:Reference number: A67369; PDB:1ZAO
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Tulinsky, A.; Mathews, I.I.
A:Reference number: A52804; PDB:1HIT
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442
R;Hrabal, R.; Komives, E.A.; Nl, F.
A:Reference number: A65583; PDB:1FGD
A:Contents: annotation; conformation by (1)H-NMR, residues 427-444
R;Hrabal, R.; Komives, E.A.; Nl, F.
Protein Sci. 5, 195-203, 1996
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the
A:Reference number: A58595; MUID:96276211
A:Contents: annotation; conformation by (1)H-NMR
C:Genetics:
A:Gene: GDB:7HBD
A:Cross-references: GDB:119613; OMTM:188040
A:Map position: 20p11.2-20p11.2
A:Introns: #status absent
C:Complex: homodimer, urinary form
C:Function:
A:Description: Inhibits thrombin activation of fibrinogen; cofactor for thrombin activat
A:Pathway: blood coagulation moderation
A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie
A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coag
e protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
F:19-513/Domain: extracellular #status predicted <EXT>
F:39-486/Product: thrombomodulin, urinary form #status experimental <MAT>
F:24-167/Domain: C-type lectin homology <LCH>
F:17-199/Region: PEST sequence
F:201-233/Region: PEST sequence
F:245-280/Domain: EGF homology <EG1>
F:288-323/Domain: EGF homology <EG2>
F:329-362/Domain: EGF homology <EG3>
F:369-404/Domain: EGF homology <EG4>
F:408-439/Domain: EGF homology <EG5>
F:445-480/Domain: EGF homology <EG6>
F:485-513/Region: PEST sequence
F:517-539/Domain: transmembrane #status predicted <TM>
F:540-575/Domain: intracellular #status predicted <INT>
F:474,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:174,225,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:252-265,267-280,288-296,292-308,310-333,329-340,336-349,351-362,369-378,374-38
F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:342/Modified site: erythro-beta-hydroxyaspartic acid (Asn) #status experimental
F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best local similarity 100.0%; Pred. No. 9e-169;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVYVIGALALAGLGFPAAPAPGPGGSCVCHDFALYPPAPFLNLSQICGIGLGHM 60
|||||
DB 1 MGVYVIGALALAGLGFPAAPAPGPGGSCVCHDFALYPPAPFLNLSQICGIGLGHM 60
|||||

QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQIPGCGDPKRLGRLFGFWVTGDNNTSYS 120
|||||

DB 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQIPGCGDPKRLGRLFGFWVTGDNNTSYS 120
QY 121 RMRLDLNGLAPLCGLPCLVAVSAEATVSEPIWEQCEKAGGFLCEFFPATCAPLAY 180
DB 121 RMRLDLNGLAPLCGLPCLVAVSAEATVSEPIWEQCEKAGGFLCEFFPATCAPLAY 180
QY 181 EPGAAAAVSTYGPFPARGADFOALPVSSAAVAPLGLCTAPPAVAGVHAREAP 240
DB 181 EPGAAAAVSTYGPFPARGADFOALPVSSAAVAPLGLCTAPPAVAGVHAREAP 240
QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGAAADGRSCATASATQSCNDLCEHCPVP 300
DB 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGAAADGRSCATASATQSCNDLCEHCPVP 300
QY 301 DQGSYSICMCEGYRLADQHRCEVDVDCILEPSPQRCVNTQGGEGECHYPPNDLVNG 360
DB 301 DQGSYSICMCEGYRLADQHRCEVDVDCILEPSPQRCVNTQGGEGECHYPPNDLVNG 360
QY 361 ECEPVPDPCEFRANCCEYOCPLNOTSYLCAEGFAPRPHRPHRCMFCNQTACPADCPN 420
DB 361 ECEPVPDPCEFRANCCEYOCPLNOTSYLCAEGFAPRPHRPHRCMFCNQTACPADCPN 420
QY 421 TQASCEPEGYILDDSFICTDIDECENGFCGVCNHLPGTFECIGPDSALVRHIGTDC 480
DB 421 TQASCEPEGYILDDSFICTDIDECENGFCGVCNHLPGTFECIGPDSALVRHIGTDC 480
QY 481 DSGKVDGSDSGSEPPSPFPSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPFPSTLTPPAVGLVHSG 516

RESULT 2
A60501
thrombomodulin precursor - mouse
N:Alternate names: fetomodulin
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence-revision 17-Apr-1993 #text-change 16-Jul-1999
C:Accession: S08488; A32001; A60501
R:Dittman, W.A.; Majerus, P.W.
Nucleic Acids Res. 17, 802, 1989
A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted
A:Reference number: S08488; MUID:89120454
A:Accession: S08488
A:Molecule type: mRNA
A:Residues: 1-577 <DIT>
A:Cross-references: EMBL:X14432; NID:954781; PIDN:CAA32597.1; PID:954782
R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.
J. Biol. Chem. 263, 15815-15822, 1988
A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat
A:Reference number: A32001; MUID:89008498
A:Accession: A32001
A:Molecule type: mRNA
A:Residues: 97-577 <DIT>
A:Cross-references: GB:J04060
R:Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.
Dev. Biol. 140, 113-122, 1990
A:Title: Identification of fetomodulin, a surface marker protein of fetal development
A:Reference number: A60501; MUID:90292331
A:Accession: A60501
A:Molecule type: protein
A:Residues: 19-22,330-343; 479-489; 545-555; 562-575 <IMA>
C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor
F:24-165/Domain: C-type lectin homology <LCH>
F:244-279/Domain: EGF homology <EG1>
F:287-322/Domain: EGF homology <EG2>
F:328-361/Domain: EGF homology <EG3>
F:368-403/Domain: EGF homology <EG4>
F:407-438/Domain: EGF homology <EG5>
F:444-479/Domain: EGF homology <EG6>

RESULT 5
MEGF6 protein - rat
Accession: F13954
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: F13954
R:Nakajima, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: 214126; MUID:98360089
A:Accession: F13954
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1578 <NAK>
A:Cross-references: EMBL:AB011532; NID:93449293; PIDN:BA32462.1; PID:93449294
A:Experimental source: strain Sprague-Dawley; brain
A:Genetics:
A:Gene: MEGF6

Query Match 14.28; Score 414; DB 2; Length 1574;
Best Local Similarity 34.18; Pred. No. 1.2e-17;
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;
QY 224 CTAPFG-AVQGMAREARAGAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA 282
DB 147 CRCPFGYQDGD-GKTCQDVDECAHNGCGHRCVMPGSLCECKPGRFLHTDRTCL- 204
QY 283 SATQSC---NDLCEHPCVP-----NPDPSGYSCK----- 309
DB 205 -AISCTYLGNGGCGHCCQVLTQVHRCQRPQYQALQEDGRRCVRRSPRCSENGSGMHIQ 263
QY 310 -----CETGYRLAADQHRCEVDVDCILEPSPORCVNTQSGFECPCPNYDL-V 359
DB 264 ELRLGHLGCHGRYQALADRKTCEDVDICALGLAAGHGLNTQSGFCKVCVHAGYELGAD 323
QY 360 G-EC-----VEPVDDCFRAN--CEYQCPRLMOTSLVCVAEGFRPIHEHRCOMC----- 408
DB 324 GRQCRIMEIVNSCEKNGGSGHSGH--TSTGRLCTCPRG-----ELDEQKTCIDID 377
QY 409 ---NOTACPADCDPNTQA--SCECPREGYILD-DGFTCTDIDECENG-GFGSCVCHNLDPGT 461
DB 378 DCANSPCCQOAC-ANTPGYEGSCFAGYRLNTDGCCEEDVDECAHSGHSGHNCNLGAS 436
QY 462 FECCIC 466
DB 437 FQCCFC 441

RESULT 6
Fibrillin-2 precursor - mouse
Accession: A57278
C:Species: Mus musculus (house mouse)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 11-Jan-2000
C:Accession: A57278
R:Zhang, H.; Hu, W.; Ramirez, F.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix
A:Reference number: A57278; MUID:95263670
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-references: GB:U39790; NID:9762830; PIDN:AAA74908.1; PID:9762831
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1239-1274/Domain: EGF homology <EGF1>
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.88; Score 373; DB 2; Length 2907;
Best Local Similarity 32.94; Pred. No. 5.7e-15;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVENGCEHACNA1PGAPRCQCPAGALQADGRSCTASATQSCNDLCEHPCVPND-- 301
DB 1238 ECMIMNGGCDTQCTNSSEGSSEGSCEGALMPDRSCA-----DIDE--CENNPDIC 1287
QY 302 -----QPGTSCCEFGYRLAADQHRCEVDVDCILEPSPC-QRCVNTQSGFECPC 353
DB 1288 DGGQCTNIGEGYRLCYGFAMSDMKTCIDVNCIDNPMICMGECENTKSGFICHCQL 1347
QY 354 NYDLVDFE--CYEPPDC--FRANCEYQCPN-OTSILVCVAEGFA-----PIR 398
DB 1348 GTSYKKGTTGCAVD-VEDEIAGHNDMHASCLANPSPKSCREGWGNKIKCIDDECA 1406
QY 399 HEHRCOMFCNCTACPADCDPRYQAS--CECPREGYILDGFTCTDIDE-----CENG- 448
DB 1407 NGTHQCSI-----NAQC-VNTPGSTRACSEGF-TGDEFTGSDVDECAENTNLCEMG 1457
QY 449 -----GF-----GVCHNLPGTFECICGPDVALY 473
DB 1458 CLNVPAYRCCECMGFTTASDSRSQDIDECSPNICEVFTCNNTLGMFHCICDDGIELD 1517
QY 474 RHIG--TDCD 481
DB 1518 RTGGNCTDID 1527

RESULT 7
Fibrillin-2 precursor - human
Accession: A55184
N:Alternate names: protein DKFZp586A1519.1
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: A55184; T08744
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
Genomics 22, 425-430, 1994
A:Title: Fibrillin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the
A:Reference number: A55184; MUID:95104855
A:Accession: A55184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1184 <ZHA>
R:Mamuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, March 1999
A:Reference number: 216471
A:Accession: T08744
A:Molecule type: mRNA
A:Residues: 656-719, QDECLMGADHCSRQRCVNTLGSFYCVNHTVLCAQGYILNHRKCV, 720-853, 'T', 855
A:Cross-references: EMBL:AL050095
A:Experimental source: adult uterus; clone DKFZp586A1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135821
A:Map position: 3p25-3p24
A:Note: DKFZp586A1519.1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; extracellular matrix
F:1-27/Domain: signal sequence #status predicted <Sig>
F:78-1184/Product: fibrillin-2 protein #status predicted <Mat>
F:905-941/Domain: EGF homology <EGF>

Query Match 12.68; Score 367.5; DB 2; Length 1184;
Best Local Similarity 30.08; Pred. No. 5.7e-15;
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;
QY 154 EEOCEVKNAD-GFLCEFFHPATC-----RPLAV-----EPGAAAAVSTYGTTPPAANG 201
DB 531 EGOSCESNPNLGPCCN-HYMLSCCEGEPFLIVPRPPPAAPRRVS-----EAM 582
QY 202 ADFQALVGGSSAAV--APLGIQLMCTAPGAVGGMAREAPAMDCSVENGCEHACNA 258
DB 583 AGREALSLGTAEALPNSLPDGDODECLLPGL-----CQHLCTIN 622

Oy	259	IPGAPROCPAGAAALADNRSC-----TASAT-----QS	287
Db	623	TVGSYHCACPGFSLDDDGTCRPGSHPRQPAROERPAKSEFSOYASNTTPLPLPQNT	682
Oy	288	CND--LCEHFCVNPPOPGSISMCESTGYRLADOHRCEDVDCILLESPPR--QRCVNT	343
Db	683	CKDNGPKQYCS--SYVGSALICSPERGALIMADGSCEDINECYTDLHTCSRGHECVNT	739
Oy	344	QGGFECH----CTPNDDLVNGECVEYVDQCFRANCEYQOQPLNQSYLCV-----	389
Db	740	LGSFHCKALTCERGALAKDGECEVDVDECE--AMGTHTOP----GFLQONTKSGSFYQA	792
Oy	390	---CAEFAPLPH-----EPHRCOMFCNOTACPADCDPNTQASCEPEGY-I	432
Db	793	RQRMRMDFLDDPESNCVDINECTSLSEPRCPGSPSCINTGSYTCQRNPLI---CARGYHA	849
Oy	433	LDDGFICTDIDECENGSGFCG---VCHNLPRGTFECIC	466
Db	850	SDDGAKCVADNCESTGVNRCGEQGVCHNLPGSTRCC	886

RESULT 8
A55567
Fibrillin 1 - bovine
C|Species: Bos primigenius taurus (cattle)
C|Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C|Accession: A55567
R|Lister, D.J.; Li, L.; Potter, K.A.; Momack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A|Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to
A|Reference number: A55567; MIM:93137597
A|Accession: A55567
A|Status: preliminary
A|Molecule type: mRNA
A|Residues: 1-2871 <TI>
A|Cross-references: GB:U28748; NID:G508427; PIDN:AA74122.1; PID:G508428
C|Superfamily: unassigned EGF-related proteins: EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match	Score 367	DB 2	Length 2871
Best Local Similarity	31.5%	Pred. NO. 1.3e-14	
Matches 111	Conservative 26	Mismatches 109	Indels 106
			Gaps 20
Qy 244	DCSVNGGCEHACNAIPGAPRCOCPPAGALQADRGSCSTASATQSCNDLCEHRCVNP	---	300
Db 1200	ECGIMNGGETFCTNSEGSTECSCGGFALMPQRST	-----	DIDE--CEDPNPIC 1249
Qy 301	-----DQPGSYSCMCETGYRLAADQRCEDVDCILLESPPC-PQRCVNTQGGFECHCYR	---	353
Db 1250	DGGCTNIPREYRCLCYDGFMASEDKTCYDVNECDLNPNICLSTGTCENTKGSFICHCM	---	1309
Qy 354	NYDLVGE--CYRVPDRC--FRANCEYQOQPLNGT--SYLCVCGEFA-----	---	PLP 398
Db 1310	GYGKRGKGTCTP--INECEIGAHNCORHAYCTTAPGSKFCSPGMIGDIKCTDIDECS	---	1368
Qy 399	HEBRCOMFCNCPACPADCPNTOAS--CEBPEGYLLDDGFTCTDIDEC--EN-----	---	447
Db 1369	NGTHMCSQH-----ADC-KNTMGSTRCLCKRGY--TGDSFTCTDDECEENLIGNGQ	---	1419
Qy 448	-----GGF-----CS-----GVCNLDPEFTFCIGPDSALY	---	473
Db 1420	CLNAPGTYRCEDMGFVPSADKACEDIDECSLPNICVFCTCHNLPGLFRCCEIEGYELD	---	1479
Qy 474	RHTG-----TDCDSGK--VDGDDSSGGEPPSPPTGSLTTPAAVLY	---	513
Db 1480	RSNGCNTDVNECDLPTTCISGNCVNTGASYTCDDPDP-----FELNPNPRVCV	---	1527

RESULT	9
A54105	
Fibrillin-2 precursor - human	

C.Species: Homo sapiens (man)
C.Date: 09-Sep-1994 #sequence,revision 09-Sep-1994 #text_change 17-Nov-2000
C.Accession: A54105, S17063; S31101
R.Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sangunetti, C.; Bonadio, J.; Mecha
J. Cell Biol. 124, 855-863, 1994
A.Title: Structure and expression of fibrillin-2, a novel microfibrillar component pr
A.Reference number: A54105; MUID:94165150
A.Accession: A54105
A.Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A.Molecule type: mRNA
A.Residues: 1-2918 <ZHA>
A.Cross-references: GB:U03272
R.Iee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
Nature 352, 330-334, 1991
A.Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
A.Reference number: S17062; MUID:91304567
A.Accession: S17063
A.Molecule type: mRNA
A.Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>
A.Cross-references: EMBL:X62009
R.Milewicz, D.M.
submitted to the EMBL Data Library, December 1992
A.Reference number: S31101
A.Accession: S31101
A.Molecule type: mRNA
A.Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P',
A.Cross-references: EMBL:X62009
C.Genetics:
A.Gene: GDB:FN2
A.Cross-references: GDB:128122; OMIM:121050
A.Map position: 5q23-5q31
C.Superfamily: unassigned EGF-related proteins; EGF homology
C.Keywords: extracellular protein
F:1-29/Domain: signal sequence #status predicted <Sig>
F:30-2918/Product: fibrillin-2 #status predicted <Mat>
F:1245-1280/Domain: EGF homology <EGF1>
F:1970-2013/Domain: EGF homology <EGF>

	Query Match	12.58;	Score 365;	DB 2;	Length 2918;	
	Best Local Similarity	33.18;	Pred. No. 1.7e-14;			
	Matches 101; Conservative	25;	Mismatches 97;	Indels 82;	Gaps 17.	
Oy	244 DCSVNGGCGCHACNAITGARPCRCQPPAGALAQADGRSCTASATPSCNDLCEHFCVPD--	301				
	: : :	:		:		
Db	1244 ECMIMNGCPTOCINSBGSEYCSCEYLAMPGRSCA-----DIDE--CENNPDIC	1293				
Oy	302 -----QPSSYSOMCEETGYRLAADQRHCEDVDVDCILEPSPC-QRCVNTQGGEFECNCP	353				
	: : :	:		:		
Db	1294 DGGCCTNIPIPGRYCLCYDGFMASDMKTCIDINECDLNSNICMGRECENTKGSGFTCHCOL	1353				
Oy	354 NYDLVDE--CVEYVDPG--FRANCEYQCQPLN-QTSYLCAEGFAPIPH-----EPHR	403				
	: : :	:		:		
Db	1354 GYSVKKGTGGTD-VDECEIGAHNCMDMHASCINIPGSFKSCREGW--IENGKICIDLDE	1410				
Oy	404 QCMCNQTACPADDPMPTQAS--CECEPYGLDDGFICTIDE-----CENG-----	448				
	: : :	:		:		
Db	1411 CSNGTHOCSIQAOC-VNTPGSTRCACSEGF-TGDGPFCSVDYDECAEAININCENGGQLNVP	1468				
Oy	449 -----GF-----CSGVCHNLPGTFEEICGPDSALVRHIG-	477				
	: : :	:		:		
Db	1469 GAYRCECEMFTPASDSRSODIDECSFONICVSGTGNLPGMHHCICDDGIELDRIGN	1528				
Oy	478 -TTCD 481					
	: :					
Db	1529 CTDID 1533					

```

RESULT 10
A47221
fibrillin 1 precursor - human (fragment)
C:Species: Homo sapiens (man)
C:date: 02-Jun-1995 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

```

C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perella, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillan, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 132-3002 <PER>
 A:Cross-references: GB:U13923; NID:9306745; PIDN:AA02036.1; PID:9306746
 R:Maalen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556
 R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831
 A:Accession: I59574
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2217-2288, 'T', 2290-2325 <RES>
 A:Cross-references: GB:S54426; NID:9264860; PIDN:AA02544.1; PID:9264861
 R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.
 Nature 352, 330-334, 1991
 A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe
 A:Reference number: S17062; MUID:91304567
 A:Accession: S17062
 A:Molecule type: mRNA
 A:Residues: 'VATVVFILSTNKM', 944-1444 <LEB1>
 A:Cross-references: EMBL:X62008; NID:913198; PIDN:CA05534.1; PID:95924015
 A:Accession: S62111
 A:Molecule type: protein
 A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEB2>
 R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21381-21385, 1989
 A:Title: Connective tissue microfibrils. Isolation and characterization of three large F
 A:Reference number: A34198; MUID:90078246
 A:Accession: A34198
 A:Molecule type: protein
 A:Residues: 565-575; 1890-1892, 'T', 1894-1900 <MAD>
 C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
 C:Genetics:
 A:Gene: FBN1
 A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
 A:Map position: 15q21.1-15q21.1
 A:Introns: 2236/1; 2258/1; 2297/1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; N
 F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted
 F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MATC>
 F:133-1367/Product: EGF homology <EGF>
 F:1457-1492/Domain: EGF homology <EGF2>
 F:2262-2295/Domain: EGF homology <EGF1>

Query Match 12.4%; Score 363; DB 2; Length 3002;
 Best Local Similarity 30.8%; Pred. No. 2, 4e-14;
 Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;
 244 DCSVNGCEHACNMPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP--- 300

```

Db      1331 ECSTIMNGCSTFCTNBSGSECCSGPFGALMPQRST-----DIE--CEDNPNC 1380
Qy      301 -----DQPSYSQMCETGYRLADQHRCEVDVDCILBSPC--PQRCVNTGGFFECPCP 353
Db      1381 DGGQCTNIPGEYRLCYDGMASEDMKTCVDFNECDLNNICISGTCENTKGSFICHCMD 1440
Qy      354 NYDLVGE--CPEPVDP--FRANCYQCPINOT--SYLCVABEFA-----PIP 398
Db      1441 GYSGKRCCKTCTD--INCEIGAHNCKHAYCTNTAGSFKCSGPGWIGDKICTDLDECS 1499
Qy      399 HEPHRCQMFQNTACPADCPNTQAS--CEPEGYILDDGFTCTDDEC--EN----- 447
Db      1500 NGTHMCSQH-----ADC--KNTMGSRCLCKREG--TGGGFTCTDDECESENLNCGNQ 1550
Qy      448 -----GGF-----CS-----GVCNLPDTECTICGPDLSLV 473
Db      1551 CLNAPGGYRCCECDMGFPASADGRACEDIDECSPNIVCGTCHNLPGLRCECEIGEYLD 1610
Qy      474 RHIG-----TDCDSGKVDGSGSGEPSPPTGSG-----TLTPPAGLV 513
Db      1611 RSGGNCCTVNECIDPTTCTISGNCVN-----TRESYICDPPPELNPTRVGCV 1658

RESULT 11
A55624
fibrillin-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000
C:Accession: A55624
R:Yin, W.; Smiley, E.; Gemmiller, J.; Sanginetti, C.; Lawton, T.; Perella, L.; Ramire
J. Biol. Chem. 270, 1798-1806, 1995
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin
A:Reference number: A55624; MUID:95130561
A:Accession: A55624
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <YIN>
A:Cross-references: GB:I29454; NID:9575509; PIDN:AAA6840.1; PID:9575510
C:Genetics:  

A:Gene: Fbn-1  

C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>
```

Query Match 12.4%; Score 361; DB 2; Length 2871;
 Best Local Similarity 29.4%; Pred. No. 3e-14;
 Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;
 212 SAAVAPRLGL-----QLMCTAPPGAVOGHWAREAPGAMDCSVENGCEHACNMPGAPRC 265
 Db 1162 SANLCPRHRCVNLIGKYQACNPGYHPTDRLFCVVIDECSIMNGCEFTCTNSDGSYEC 1221
Qy 266 QCPAGALQADGRSCTASATQSCNDLCEHFCVNP-----DQPSYSQMCETGYRL 316
Db 1222 SCQPFALMPDRSCT-----DIDQ--CEDNPNICDGGQCTNIPGEYRLCYDGFMA 1271
Qy 317 AADHRCEDVDCTLEBSPC--PQRCVNTGGFFCHCPYNDLVGE--CPEVPDP--PR 371
Db 1272 SEDMKTCVVDVNECDLNNITLSTGTCENTKGSFICHDMSGKKGCTCTD--INCEIETA 1330
Qy 372 ANCEYQCPINOT--SYLCVABEFA-----PIPHPRCQMFQNTACPADCPDN 420
Db 1331 HNCGRHNVGNTAGSFKCSGPGWIGDKICTDLDECSNGTHMCSQH-----ADC--KN 1382
Qy 421 TQAS--CEPEGYILDDGFTCTDDEC--EN-----GGF----- 450
Db 1383 TMSYRCLCKDXY--TGDGFTCTDDECESENLNCGNQCLNAPGGYRCCECDMGFPVSADG 1441
Qy 451 -----CS-----GVCNLPDTECTICGPDLSLVHIG-----TDCDSGK 484
Db 1442 KACEDIDECSPNIVCGTCHNLPGLRCECEIGELDRSGGNCCTVNECIDPTTCTISGN 1501

OY 485 VDGSDSGSGPPSPSPGS-----TLPPAVGLV 513
 DB 1502 CVN-----TPGSTYCCDCSPDFELNPRVGCY 1527

RESULT 12

46488
 hypothetical protein DKFZp434J065.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46488

R:Diesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035

A:Accession: T46488

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-741 <AAA>

A:Cross-references: EMBL:AL137638

A:Experimental source: adult testis; clone DKFZp434J065

C:Genetics:

A:Note: DKFZp434J065.1

Query Match

Best Local Similarity 12.3%; Score 359; DB 2; Length 741;
 Pred. No. 1.2e-14;

Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;

OY 245 CSVNGCENACNAIPGAPRCQCPAGALADGRSCTA-SATQSCNDLCHEFCVPPNDP 303

DB 68 CAMBDHNCCEOLCVNPPSCYQCCYSGYALAEGRKCAVAVYCASENHGCHECV-NAD-- 124

OY 304 GSYSCMCEGYRLAADHRCEDVDVDCILEPSPCPCRCVNTQGGFECVPPNDP-VNGEC 362

DB 125 GSYLCQCHGEFALNPDECTCKIDYICASSNHGCOHECVNTDDSYSCGLGFTLNPDKT 184

OY 363 VEPVDPCE--RANCEYQCPNLQTSYLCVCAEGFAPRPH-----EPHRCQMC 408

DB 185 CRRINCYCALMKPCGEHNECVNME--SYVCRHRYGTYLDPNGKTCRVNHCADQDHGCEBOLC 243

OY 409 NOTACPACDNDPTQAS--CECPGTYIIDDF--ICTDIDEC--ENGFGSGVCHNLPGTF 462

DB 244 -----LNTEDSFYCCQCSGFLINEDLKTCSRDVYCLLSDHG--CERYCVNMDRSF 291

OY 463 ECICGPDALVRHIGTDDSGKVDG---GDSG 491

DB 292 ACQC-PEGHVLRSQDKTC--AKLDSICALGDHG 320

RESULT 13

449457
 fibulin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000

C:Accession: A49457; S74095

R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.

J. Cell Biol. 123, 1269-1277, 1993

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with

A:Reference number: A49457; MUID:94064787

A:Accession: A49457

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1221 <PAN>

A:Cross-references: GB:X75285; NID:9437046; PIDN:CAA53040.1; PID:9437047

R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.

Eur. J. Biochem. 240, 427-434, 1996

A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met

A:Reference number: S74094; MUID:96439073

A:Accession: S74095

A:Molecule type: protein

A:Residues: 236-238, 'X', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-589

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer

F:942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;
 Best Local Similarity 29.5%; Pred. No. 2.5e-14;
 Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

OY 154 EEOCEVKAD-GELEFHPATC-----RPLAV-----EFGAAVAAS-----IT 192

DB 521 EEOCEBSPNLTGYPCH-HVMSLSCGEPLVPEVRRPEBPAPRVSEMAAREALS 579

OY 193 YGT---PFAAGADFO---ALP-----VGSSAAVAPLQLM-----CTAPPG 229

DB 580 LGTEALPNSLPGDDODECLMFGELCHLCTINTVGYRCAFCPEFELGGGRICRPDRG 639

OY 230 AVQGHARE-ARGANDCSY-----ENGCEHACNAITGAPRCQCPAGAA 272

DB 640 APQLDTAREASAPRSASQVSPNTIPLPVQPNPTCKDNPCRQVCRVVDGTAMCSCFPGYA 699

OY 273 LDADGRSC-----TASAFQSCNDLCHEFCVPPNDPQGSYSC-----MCEFGYRLAADHR 322

DB 700 IMADGYSCEDDOECLMGTHDCS--WKQFCV---NTLGSFYCVNHTVLCABEYILNA-HRK 753

OY 323 CEDVDVDCILEPSPC--PQRCVNTQGGFECV---CYPNYDLVDGCEVPPVDPCEFA--NC 374

DB 754 CVDINCEVTDLHTCTRAEHCVNTPGSCYKALTCPEGVLTGBCSD-VDECVGTGHC 812

OY 375 E--YQCPNLQTSYLCV---CAGEFAPRPH-----EPHRCQMCNQTACPA 415

DB 813 QAGFSCQ--TKGSFYCQARQCMDFLDPREGNCVDINECTSLLEPCSSGFCINTVGSY 871

OY 416 DCDPNTQASCEPREG-IIDDFICTDIDECENGFGCSG---VCHNLPGTECTICGP 468

DB 872 TQORNPV---GGRGYHANEBSSECVDVNECTGVHNRGEGQLCVNLGYSRCDCKP 925

RESULT 14
 743210
 fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and charac

A:Reference number: 223337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match

Best Local Similarity 11.4%; Score 331; DB 2; Length 589;
 Pred. No. 5e-13;

Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

OY 103 GPLRGF-QWVTGPNNTSYRMRARLDNGAPRLCGPLCVASAAEAVPSPRI-----WEQ 156

DB 19 GCLRSFNKCCNGDEITLH--ASEIITGRLNDPVLHLDRCASSHCHLCHDRGSKV 75

OY 157 QCEVKA-----DGFICEFHP--ATCRPLAVE-----PGAAAASITVGPFA-- 198

DB 76 EECRCRSGFLDAPGMACVNHIDECATLMDDCLESQRCRLTTPSPFCIRLSGCTGYAADS 135

OY 199 --ARGADFQALPYGSSAAVAAPLGLDLMCTAPPAVQ-----GHMAREAPGAMDCS-- 246

DB 136 ETERCRDVEDCNLGSN---DCGPLYOCRNTOGSYRCAKCKGDELDLNPMTGECTSTTC 191

OY 247 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGALADGRSCTA 282

Db	192	PNGYPRKMGMCNDIDECVTGHCNCGAEECVNTPGSGFOQKGNLCAGHYEVN-----	243
Qy	283	SATQSCNDL--CEH-----PCVNPDPGSGYSOMCEGYSGLADNRHCEVDPCIL---	331
Db	244	GATGFCBEVYNNCCQGVGSGMECT---NLPGYRKCKCGGYGFNNAKRKCEVEDEICIFAG	300
Qy	332	EPSPQPCRCVNTGGGFECHCTPNTDLV-DGECVEPVDCEFR--ANCEYQCPNLQSTSLC	388
Db	301	HVCDLSAECINTIGSFCECKRPGQLASDGRRCADYNECTGIAGCQKCVNI--PGSQOC	359
Qy	389	VCAGSFAPIP-----HEPHRCOMFC--NOTACPADCDPNTQAS--CECPGEGY-ILDDGFI	438
Db	360	ICDRGFALGPDTGCKEIDEDICSIVANGSNDLCWGGC-INTKGYSLQCPPGYKIQPDGRT	418
Qy	439	CTDIDECENSGFCSC---YCHNLPGTFEC	464
Db	419	CVDVDECCAMG-CAGSDKRCVNTLGSFRC	446

OY	451	C-----	SGCHNLPTETFCIGP-----	448
Db	359	CONFAGSFHCVCASGWSGAGCEENLMDCAATCAPGSCYCIDRVGSFSCILPPGRGTLLCH		368
OY	469	-DSALVR-HIGTDCDSGRVDS-----	GDGGS-----	459
Db	369	LEDMLSPCHVHNOCCSTNPLTGSTLCIGPGYSGSCTHODLDECCMAQGGSPSCHEHGS		448
OY	500	---TPGS 503		
Db	449	CINTPGS 455		

Search completed: May 8, 2002, 12:40:00
Job time: 246 sec

RESULT 15
T09059
notch4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jan-2000
C:Accession: T09059
R:Rowen, J.; Mahatiras, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Schmitt, J.
Submitted to the EMBL Data Library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: 216543
A:Accession: T09059
A:Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1964 <ROW>
A:Cross-references: EMBL:AF030001; NID:g2564945; PTD:g2564947
C:Genetics:
A:Gene: notch4
A:Map position: 17
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 671/1
1679/3; 1729/1; 1761/3
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: receptor; signal transduction
;514-545/Domain: EGF homology <EGF>

Query Match	11.3%	Score 329.5;	DB 2,	Length 1964;
Fest Local Similarity	26.3%	Pred. No. 1.7e-12;		
Matches 128; Conservative	33;	Mismatches 143;	Indels 183;	Gaps 31;

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:39:20 ; Search time 21.32 seconds
(without alignments)
544,639 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MLGVVLGALAGLGPAP.....PSPTPGSLTPPAGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2916	100.0	575	1	US-08-312-870-1
2	2916	100.0	575	6	Sequence 1, Appli Patent No. 5466668
3	2912	99.9	575	1	US-08-261-206A-59
4	2908	99.7	575	1	Sequence 59, Appli Patent No. 5256770
5	2846.5	97.6	572	6	Sequence 54, Appli Patent No. 5256770
6	2826	96.9	498	2	Sequence 2, Appli Patent No. 5256770
7	2824	96.8	497	1	Sequence 3, Appli Patent No. 5256770
8	2766	94.9	494	1	Sequence 14, Appli Patent No. 5256770
9	2766	94.9	494	1	Sequence 14, Appli Patent No. 5256770
10	2764	94.8	494	1	Sequence 14, Appli Patent No. 5256770
11	2764	94.8	494	1	Sequence 14, Appli Patent No. 5256770
12	2690	92.2	475	1	US-08-110-011A-16
13	2690	92.2	475	1	US-08-307-444A-2
14	2686	92.1	475	1	US-08-587-389-2
15	2686	92.1	475	1	US-08-307-444A-1
16	2680	91.9	476	1	US-08-587-389-1
17	2680	91.9	476	1	US-08-014-723-1
18	2678	91.8	476	1	US-08-110-011A-1
19	2678	91.8	476	1	US-08-014-723-2
20	2678	91.8	476	1	US-08-014-723-18
21	2678	91.8	476	1	US-08-110-011A-2
22	2596	89.0	456	1	US-08-110-011A-18
23	2596	89.0	456	1	US-08-307-444A-4
24	2592	88.9	456	1	US-08-587-389-4
25	2592	88.9	456	1	US-08-307-444A-3
26	2543	87.2	446	1	US-08-587-389-3
27	2543	87.2	446	1	US-08-307-444A-5
					Sequence 5, Appli Patent No. 5466668

28	1621	55.6	275	1	US-08-312-870-7	Sequence 7, Appli
29	1159	39.7	215	1	US-08-312-870-5	Sequence 5, Appli
30	689	23.6	115	1	US-08-312-870-9	Sequence 9, Appli
31	681	23.4	114	2	US-08-733-564-1	Sequence 1, Appli
32	584	20.0	652	2	US-08-751-305-2	Sequence 2, Appli
33	356	12.2	638	2	US-08-897-443-1	Sequence 1, Appli
34	352	12.1	58	1	US-08-261-206A-3	Sequence 3, Appli
35	326.5	11.2	956	2	US-08-897-443-3	Sequence 3, Appli
36	323	11.1	1253	3	US-08-479-722B-4	Sequence 4, Appli
37	322.5	11.1	1394	6	5177197-30	Patent No. 5177197
38	319.5	11.0	1833	3	US-08-479-722B-2	Sequence 2, Appli
39	319.5	11.0	1833	5	PCT-US95-02251-18	Sequence 18, Appli
40	315.5	10.8	443	2	US-08-833-963C-2	Sequence 2, Appli
41	315.5	10.8	443	3	US-08-980-514-1	Sequence 1, Appli
42	308	10.6	448	2	US-08-884-072-1	Sequence 1, Appli
43	308	10.6	448	4	US-09-212-168-1	Sequence 1, Appli
44	303.5	10.4	1251	5	PCT-US95-02251-3	Sequence 3, Appli
45	303.5	10.4	1252	1	US-08-199-780-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 6.9e-199;

[illegible]

RESULT 2
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,346
; FILING DATE: 22-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 568,456
; FILING DATE: 15-AUG-1990
; APPLICATION NUMBER: 506,325
; FILING DATE: 09-APR-1990
; APPLICATION NUMBER: 406,941
; FILING DATE: 13-SEP-1989
; APPLICATION NUMBER: 345,374
; FILING DATE: 28-APR-1989
; SEQ ID NO: 6
; LENGTH: 575
5466668-6

Db	61	TVRSSVAADYISLLNGDGGVRRRLWIGLQPPGGCDPKRLGPGFQWVTGDNNTSYS	120
QY	121	RWARLDLNGAPLCGPICVAVSAAEATVPSEPTWEEQOCEVKADGFLCEPHFPATCRPLAV	180
Db	121	RWARLDLNGAPLCGPICVAVSAAEATVPSEPTWEEQOCEVKADGFLCEPHFPATCRPLAV	180
QY	181	EPGAAAAVSIITYGTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPAGAVOGHWAREAP	240
Db	181	EPGAAAAVSIITYGTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPAGAVOGHWAREAP	240
QY	241	GAWDCSVENGGBHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVCPNP	300
Db	241	GAWDCSVENGGBHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVCPNP	300
QY	301	DOFGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
Db	301	DOFGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
QY	361	ECVEPVDPFRANCEYQCPNLNTSYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN	420
Db	361	ECVEPVDPFRANCEYQCPNLNTSYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN	420
QY	421	TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPPDSALARHIGTDC	480
Db	421	TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPPDSALARHIGTDC	480
QY	481	DSGKVGDDSGSGEPPPPSTLTPPAVGLVHSG	516
Db	481	DSGKVGDDSGSGEPPPPSTLTPPAVGLVHSG	516

```

3
RESULT
US-08-261-206A-59
; Sequence 59, Application US/08261206A
; Patent NO. 5574007
; GENERAL INFORMATION:
; APPLICANT: Zushi, Mitichitaka
; APPLICANT: Gomi, Komakazu
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Suzuki, Koji
; APPLICANT: Matsuda, Akio
; TITLE OF INVENTION: A Polypeptide Capable of Interacting
; TITLE OF INVENTION: with Thrombin
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,206A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/740,492
; FILING DATE: 03-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30330
; REFERENCE/DOCKET NUMBER: 216-275P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 59:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..575
OTHER INFORMATION: /label= protein
OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match 99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.3e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVGLALAGLGFAPAPAEPOGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVGLALAGLGFAPAPAEPOGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240

QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300
DB 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300

QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYPNYDLVDG 360

QY 361 ECVEPVPDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGFCGCGVCHNLPGTFCICGPDALSARHIGTDC 480
DB 421 TQASCECPGYILDDGFICTDIDECENGFCGCGVCHNLPGTFCICGPDALSARHIGTDC 480

QY 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 4
US-08-170-290A-54
Sequence 54, Application US/08170290A
Patent No. 5702931
GENERAL INFORMATION:
APPLICANT: Andrews, William H.
APPLICANT: Morser, Michael J.
APPLICANT: Ziehlender, Laura R.
TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
TITLE OF INVENTION: Compositions
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: James M. Heslin
STREET: 379 Lytton Ave.
CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/170,290A
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05573
FILING DATE: 01-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/724,237
FILING DATE: 01-JUL-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 11972-58-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-170-290A-54

Query Match 99.7%; Score 2908; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.6e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVGLALAGLGFAPAPAEPOGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVGLALAGLGFAPAPAEPOGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240

QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300
DB 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300

QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYPNYDLVDG 360

QY 361 ECVEPVPDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGFCGCGVCHNLPGTFCICGPDALSARHIGTDC 480
DB 421 TQASCECPGYILDDGFICTDIDECENGFCGCGVCHNLPGTFCICGPDALSARHIGTDC 480

QY 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

Db 481 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 516

RESULT 5

5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7
; LENGTH: 572
5256770-7

Query Match 97.6%; Score 2846.5; DB 6; Length 572;
Best Local Similarity 98.6%; Pred. No. 5.7e-194;
Matches 509; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 1 MGVVLGALAGLGGFPAPBPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
|||||
Db 1 MGVVLGALAGLGGFPAPBPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
|||||
QY 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPCGDPRKRLGFGQWVTGDNNTSYS 120
|||||
Db 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPCGDPRKRLGFGQWVTGDNNTSYS 119
|||||
QY 121 RWRDLNGLAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAV 180
|||||
Db 120 RWRDLNGLAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAV 179
|||||
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAP 240
|||||
Db 180 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTA--GNVQGHWAREAP 237
|||||
QY 241 GAWDCSVENGGEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||
Db 238 GAWDCSVENGGEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 297
|||||
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCPNYDLVDG 360
|||||
Db 298 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCPNYDLVDG 357
|||||
QY 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
|||||
Db 358 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 417
|||||
QY 421 TQASCEPGEYLLDDGFTCTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 480
|||||
Db 418 TQASCEPGEYLLDDGFTCTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 477
|||||
QY 481 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 516
|||||
Db 478 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 513
|||||

RESULT 6

US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia

; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-733-564-2

Query Match 96.9%; Score 2826; DB 2; Length 498;

Best Local Similarity 99.8%; Pred. No. 1.4e-192;

Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 78
Db 1 APAEPQPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 60
|||||
QY 79 GGVGRRRLWIGLQLPPGCGDPRKRLGFGQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 138
|||||
Db 61 GGVGRRRLWIGLQLPPGCGDPRKRLGFGQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 120
|||||
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAVEPGAAAASVITYGTFFA 198
|||||
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAVEPGAAAASVITYGTFFA 180
|||||
QY 199 ARGADFQALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGGEHACNA 258
Db 181 ARGADFQALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGGEHACNA 240
|||||
QY 259 IPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318
Db 241 IPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
|||||
QY 319 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCPNYDLVDCGECVPEVPDPCFRANCEYQC 378
Db 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCPNYDLVDCGECVPEVPDPCFRANCEYQC 360
|||||
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPNQTASCEPGEYLLDDGFI 438
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPNQTASCEPGEYLLDDGFI 420
|||||
QY 439 CTDDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSDSGEPPPPS 498
Db 421 CTDDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSDSGEPPPPS 480
|||||
QY 499 PTPGSTLTPPAVGLVHSG 516
Db 481 PTPGSTLTPPAVGLVHSG 498
|||||

RESULT 7

US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625

GENERAL INFORMATION:
APPLICANT: Carson, Craig W.
APPLICANT: Esmon, Charles T.
TITLE OF INVENTION: Method for Detecting Antibodies to
TITLE OF INVENTION: Thrombomodulin in Patients
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richards, Medlock & Andrews
STREET: 1201 Elm Street, Suite 4500
CITY: Dallas
STATE: Texas
COUNTRY: US
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,870
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hansen, Eugenia S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: OMRF B35150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-939-4500
TELEFAX: 214-939-4600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-870-3

Query Match 96.8%; Score 2824; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.9e-192;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRGLMTVRSVAADVISLLNGD 78
|||||
Db 1 APAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRGLMTVRSVAADVISLLNGD 60
|||||

QY 79 GGVGRRRLIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCGPLCV 138
|||||
Db 61 GGVGRRRLIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCGPLCV 120
|||||

QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEFGAAAAVSTYGTFFA 198
|||||
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEFGAAAAVSTYGTFFA 180
|||||

QY 199 ARGADFOALPVGSSAAVAPLGQLMCTAPPVAVQGHWAREAPAGWDCSVENGCEHACNA 258
|||||
Db 181 ARGADFOALPVGSSAAVAPLGQLMCTAPPVAVQGHWAREAPAGWDCSVENGCEHACNA 240
|||||

QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQSGYSKCMETGYRLAA 318
|||||
Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQSGYSKCMETGYRLAA 300
|||||

QY 319 DQHRCEDDVDCILEPSPCQPCQVNTGGFECYCNVLDVGGCEVPDPCFRANCEYQC 378
|||||
Db 301 DQHRCEDDVDCILEPSPCQPCQVNTGGFECYCNVLDVGGCEVPDPCFRANCEYQC 360
|||||

QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEGYILDDGFI 438
|||||
Db 361 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEGYILDDGFI 420
|||||

QY 439 CTTDIDCENGGFCGVCNHLPGTFECICGPDSSALARIHIGTDCDSKGVGDGSGSGEPPPS 498
|||||
Db 421 CTTDIDCENGGFCGVCNHLPGTFECICGPDSSALARIHIGTDCDSKGVGDGSGSGEPPPS 480
|||||

QY 499 PTPGSTLTTPPAVGLVHS 515
|||||
Db 481 PTPGSTLTTPPAVGLVHS 497
|||||

RESULT 8
US-08-014-723-14
Sequence 14, Application US/08014723
Patent No. 5273962
GENERAL INFORMATION:
APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Ohkuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,723
FILING DATE: 19930208
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5273962man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-071-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-014-723-14

Query Match 94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGFPAEPQGGSCQVEHDFCFALYPCGATFLNASQICDGLRGLH 60
|||||
Db 1 MLGVVLGALALAGLGFPAEPQGGSCQVEHDFCFALYPCGATFLNASQICDGLRGLH 60
|||||

QY 61 TVRSSVAADVISLLNGDGGVRRRLIGLQPPCGDPRKRLGRLGFQWVTGDNNTSY 120
|||||
Db 61 TVRSSVAADVISLLNGDGGVRRRLIGLQPPCGDPRKRLGRLGFQWVTGDNNTSY 120
|||||

QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
|||||
Db 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
|||||

QY 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGQLMCTAPPVAVQGHWAREAP 240
|||||
Db 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGQLMCTAPPVAVQGHWAREAP 240
|||||

QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||

Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDSSALAHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 9
US-08-110-011A-14
; Sequence 14, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-14

Query Match 94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALAGLGFAPAEPPQGGSCQVHDCFALYPGPATFLNASQICDGLRGLHM 60
Db 1 MLGVLVLGALAGLGFAPAEPPQGGSCQVHDCFALYPGPATFLNASQICDGLRGLHM 60

QY 61 TYRSSVAADVISLLNGDGVGRRRLWIGLQPLPPGCGDKRLGRLGFQVWTGDNNTSYS 120
Db 61 TYRSSVAADVISLLNGDGVGRRRLWIGLQPLPPGCGDKRLGRLGFQVWTGDNNTSYS 120
QY 121 RWARDLNGAPLCGPLICVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAY 180
Db 121 RWARDLNGAPLCGPLICVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAY 180
QY 181 EPGAAAAVITYGTPFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
Db 181 EPGAAAAVITYGTPFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDSSALAHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 10
US-08-014-723-16
; Sequence 16, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELETYPE: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:

;
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-014-723-16

Query Match 94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 3.3e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAPAEPPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFAPAPAEPPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWRDLNLCAPLCGPLCVAVSAEAATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
DB 121 RWRDLNLCAPLCGPLCVAVSAEAATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
DB 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPCPCQRCVNTQGGFECHCYPNYDLVDG 360
DB 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPCPCQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
DB 361 ECVEPVDFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCICGPDPSALVRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSD 494

RESULT 11
US-08-110-011A-16
; Sequence 16, Application US/08110011A
; Patent No. 5354564
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-16

Query Match 94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 3.3e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAPAEPPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFAPAPAEPPGSGSQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWRDLNLCAPLCGPLCVAVSAEAATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
DB 121 RWRDLNLCAPLCGPLCVAVSAEAATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
DB 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPCPCQRCVNTQGGFECHCYPNYDLVDG 360
DB 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPCPCQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
DB 361 ECVEPVDFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCICGPDPSALVRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSD 494

RESULT 12
US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307.444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-307-444A-2

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Qy 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307.444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-307-444A-2

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Qy 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCQAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCEHACNAIPGAPRCQCQAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DQPGSYSCMCEYGRVLAADQHRCEVDVDCILEPSPCQRCQVNTQGGFECCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCEYGRVLAADQHRCEVDVDCILEPSPCQRCQVNTQGGFECCHCYPNYDLVDG 360
Qy 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACFADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACFADCDPN 420
Qy 421 TOASCCEPGEYILDDGFICTDIDECENGFGCSGVCHNLPGTFCICGPDPSALARH 475
Db 421 TOASCCEPGEYILDDGFICTDIDECENGFGCSGVCHNLPGTFCICGPDPSALARH 475
```

```
Db 421 TOASCCEPGEYILDDGFICTDIDECENGFGCSGVCHNLPGTFCICGPDPSALARH 475

RESULT 13
US-08-587-389-2
; Sequence 2, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-587-389-2

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLGFGQWVTGDNNTSYS 120
Qy 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARLDNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
```

QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCECEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALARH 475
Db 421 TQASCECEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALARH 475
RESULT 14
US-08-307-444A-1
; Sequence 1, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-307-444A-1
Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVLGALAGLGFAPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

Db 1 MLGVLVLGALAGLGFAPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TYRVSSVAADVLSLLNGDGGVGRRLWIGLQLPFGCDPKRLGRLGFQVNTGDNNTSYS 120
Db 61 TYRVSSVAADVLSLLNGDGGVGRRLWIGLQLPFGCDPKRLGRLGFQVNTGDNNTSYS 120
QY 121 RWARLDNLGAPLCGCLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 180
Db 121 RWARLDNLGAPLCGCLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 180
QY 181 EPGAAAAYVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAYVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCECEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALARH 475
Db 421 TQASCECEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRH 475
RESULT 15
US-08-587-389-1
; Sequence 1, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-587-389-1

Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MLGVLVLGALALAGLGFAPAEPPGGSCQVEHDCFALYPGPFATFLNASQICDGLRGHLM	60
Db	1	MLGVLVLGALALAGLGFAPAEPPGGSCQVEHDCFALYPGPFATFLNASQICDGLRGHLM	60
Qy	61	TVRSSAADVISLLNGDGGVRRRLWTGLQPPGCGDPKRLGPLRGFQWVTGDNNTSYS	120
Db	61	TVRSSAADVISLLNGDGGVRRRLWTGLQPPGCGDPKRLGPLRGFQWVTGDNNTSYS	120
Qy	121	RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV	180
Db	121	RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV	180
Qy	181	EPGAAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Db	181	EPGAAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Qy	241	GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP	300
Db	241	GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP	300
Qy	301	DOPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
Db	301	DOPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
Qy	361	ECVEPVDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN	420
Db	361	ECVEPVDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN	420
Qy	421	TQASCEPGEYILDDGFICTDIDECENGCGCGVCHNLPGTFECICGPDLSALAH	475
Db	421	TQASCEPGEYILDDGFICTDIDECENGCGCGVCHNLPGTFECICGPDLSALVRH	475

Search completed: May 8, 2002, 12:39:21
Job time: 212 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:49 ; Search time 21.32 Seconds
(without alignments)
544.639 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGVVLGALAGLGPAP.....PSPTPGSLTPPAVCLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	575	1 US-08-261-206A-59	Sequence 59, Appl
2	2912	99.9	575	1 US-08-312-870-1	Sequence 1, Appl
3	2912	99.9	575	6 5466668-6	Patent No. 5466668
4	2904	99.6	575	1 US-08-170-290A-54	Sequence 54, Appl
5	2842.5	97.5	572	6 5256770-7	Patent No. 5256770
6	2830	97.1	498	2 US-08-733-564-2	Sequence 2, Appl
7	2830	96.7	497	1 US-08-312-870-3	Sequence 3, Appl
8	2770	95.0	494	1 US-08-014-723-14	Sequence 14, Appl
9	2770	95.0	494	1 US-08-110-011A-14	Sequence 14, Appl
10	2768	94.9	494	1 US-08-014-723-16	Sequence 16, Appl
11	2768	94.9	494	1 US-08-110-011A-16	Sequence 16, Appl
12	2690	92.2	475	1 US-08-307-444A-1	Sequence 1, Appl
13	2690	92.2	475	1 US-08-587-389-1	Sequence 1, Appl
14	2686	92.1	475	1 US-08-307-444A-2	Sequence 2, Appl
15	2686	92.1	475	1 US-08-587-389-2	Sequence 2, Appl
16	2684	92.0	476	1 US-08-014-723-1	Sequence 1, Appl
17	2684	92.0	476	1 US-08-110-011A-1	Sequence 1, Appl
18	2682	92.0	476	1 US-08-014-723-2	Sequence 2, Appl
19	2682	92.0	476	1 US-08-014-723-18	Sequence 18, Appl
20	2682	92.0	476	1 US-08-110-011A-2	Sequence 2, Appl
21	2682	92.0	476	1 US-08-110-011A-18	Sequence 18, Appl
22	2596	89.0	456	1 US-08-307-444A-3	Sequence 3, Appl
23	2596	89.0	456	1 US-08-587-389-3	Sequence 3, Appl
24	2592	88.9	456	1 US-08-307-444A-4	Sequence 4, Appl
25	2592	88.9	456	1 US-08-587-389-4	Sequence 4, Appl
26	2543	87.2	446	1 US-08-307-444A-5	Sequence 5, Appl
27	2543	87.2	446	1 US-08-587-389-5	Sequence 5, Appl

28	1617	55.5	275	1	US-08-312-870-7	Sequence 7, Appl
29	1159	39.7	215	1	US-08-312-870-5	Sequence 5, Appl
30	685	23.5	114	2	US-08-733-564-1	Sequence 1, Appl
31	685	23.5	115	1	US-08-312-870-9	Sequence 9, Appl
32	580	19.9	652	2	US-08-751-305-2	Sequence 2, Appl
33	358	12.3	638	2	US-08-897-443-1	Sequence 1, Appl
34	352	12.1	58	1	US-08-261-206A-3	Sequence 3, Appl
35	328.5	11.3	956	2	US-08-897-443-3	Sequence 3, Appl
36	323	11.1	1253	3	US-08-479-722B-4	Sequence 4, Appl
37	320	11.0	1394	6	5177197-30	Patent No. 5177197
38	319.5	11.0	1833	3	US-08-479-722B-2	Sequence 2, Appl
39	319.5	11.0	1833	5	PCT-US95-02251-18	Sequence 18, Appl
40	315.5	10.8	443	2	US-08-833-963C-2	Sequence 2, Appl
41	315.5	10.8	443	3	US-08-980-514-1	Sequence 1, Appl
42	308	10.6	448	2	US-08-884-072-1	Sequence 1, Appl
43	308	10.6	448	4	US-09-212-168-1	Sequence 1, Appl
44	303.5	10.4	1251	5	PCT-US95-02251-3	Sequence 3, Appl
45	303.5	10.4	1252	1	US-08-199-780-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-261-206A-59
; Sequence 59, Application US/08261206A,
; Patent No. 5574007
; GENERAL INFORMATION:
; APPLICANT: Zushi, Mitichitaka
; APPLICANT: Gomi, Komakazu
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Suzuki, Koji
; APPLICANT: Matsuda, Akio
; TITLE OF INVENTION: A Polypeptide Capable of Interacting
; TITLE OF INVENTION: with Thrombin
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,206A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/740,492
; FILING DATE: 03-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30330
; REFERENCE/DOCKET NUMBER: 216-275P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 1e-198;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFAPAPPEQGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVLGALALAGLGFAPAPPEQGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGRLGFWQVWGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGRLGFWQVWGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180

Qy 181 EPGAAAASVTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240

Qy 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGSCSTASQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGSCSTASQSCNDLCEHFCVNP 300

Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360

Qy 361 ECVEPVPDFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVPDFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECEGYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
Db 421 TQASCECEGYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480

Qy 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carlson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match 99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.9e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFAPAPPEQGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVLGALALAGLGFAPAPPEQGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGRLGFWQVWGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGRLGFWQVWGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180

Qy 181 EPGAAAASVTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240

Qy 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGSCSTASQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGSCSTASQSCNDLCEHFCVNP 300

Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360

Qy 361 ECVEPVPDFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVPDFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECEGYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
Db 421 TQASCECEGYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480

Qy 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 3
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/155,346
FILING DATE: 22-NOV-1993
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: 568,456
FILING DATE: 15-AUG-1990
APPLICATION NUMBER: 506,325
FILING DATE: 09-APR-1990
APPLICATION NUMBER: 406,941
FILING DATE: 13-SEP-1989
APPLICATION NUMBER: 345,374
FILING DATE: 28-APR-1989
SEQ ID NO: 6
LENGTH: 575
546668-6

Query Match 99.9%; Score 2912; DB 6; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.9e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120

Qy 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180

Qy 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300

Qy 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECYPNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECYPNYDLVDG 360

Qy 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420

Qy 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
Db 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 4
US-08-170-290A-54
Sequence 54, Application US/08170290A
Patent No. 5702931
GENERAL INFORMATION:
APPLICANT: Andrews, William H.
APPLICANT: Morser, Michael J.
APPLICANT: Ziehlender, Laura R.
TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
TITLE OF INVENTION: Compositions
NUMBER OF INVENTIONS: 63
CORRESPONDENCE ADDRESS:
ADDRESS: James M. Heslin
STREET: 379 Lytton Ave.
CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/170,290A
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05573
FILING DATE: 01-JUL-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/724,237
FILING DATE: 01-JUL-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 11972-58-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-170-290A-54

Query Match 99.6%; Score 2904; DB 1; Length 575;
Best Local Similarity 99.6%; Pred. No. 7.1e-198;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120

Qy 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180

Qy 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300

Qy 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECYPNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECYPNYDLVDG 360

Qy 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420

Qy 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
Db 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

Db 481 DSGKVDGSDSGEPPSPPTGSLTTPAVGLVHSG 516

RESULT 5
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7:
; LENGTH: 572
5256770-7

Query Match 97.58; Score 2842.5; DB 6; Length 572;
Best Local Similarity 98.4%; Pred. No. 1.6e-193;
Matches 508; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
QY 1 MGVVLGALAGLGPAPAPGPGSQCVHDFCFALYPGPFLNASQICDGLRGLHM 60
DB 1 MGVVLGALAGLGPAPAPGPGSQCVHDFCFALYPGPFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPLPGCGDKPKRLGFWGTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPLPGCGDKPKRLGFWGTGDNNTSYS 119
QY 121 RWARLDLNGAPLCPGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAV 180
DB 120 RWARLDLNGAPLCPGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAV 179
QY 181 EPGAAAASVTYGTPTFAARGADQFALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAP 240
DB 180 EPGAAAASVTYGTPTFAARGADQFALPVGSSAAVAPLGLQLMCTA--GNVQGHWAREAP 237
QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 238 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 297
QY 301 DQPGSYSCMCTGYRLAADQHRCEBDVDCILEPSPQRCVNTQGGFECVNPYLDVG 360
DB 298 DQPGSYSCMCTGYRLAADQHRCEBDVDCILEPSPQRCVNTQGGFECVNPYLDVG 357
QY 361 ECVEPVDPCEFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPN 420
DB 358 ECVEPVDPCEFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPN 417
QY 421 TQASCEPEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480
DB 418 TQASCEPEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 477
QY 481 DSGKVDGSDSGEPPSPPTGSLTTPAVGLVHSG 516
DB 478 DSGKVDGSDSGEPPSPPTGSLTTPAVGLVHSG 513

RESULT 6
US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia

; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-733-564-2

Query Match 97.18; Score 2830; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 1e-192;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 APAPQPQGGSCVHDFCFALYPGPFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78
DB 1 APAPQPQGGSCVHDFCFALYPGPFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60
QY 79 GGVRRRLWTGLQPLPGCGDKPKRLGFWGTGDNNTSYSRWARLDLNGAPLCPGLCV 138
DB 61 GGVRRRLWTGLQPLPGCGDKPKRLGFWGTGDNNTSYSRWARLDLNGAPLCPGLCV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSIYGTGPPA 198
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSIYGTGPPA 180
QY 199 ARGADQFALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGCEHACNA 258
DB 181 ARGADQFALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGCEHACNA 240
QY 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318
DB 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
QY 319 DQHRCEBDVDCILEPSPQRCVNTQGGFECVNPYLDVDCVCEVDPDFRANCEYQC 378
DB 301 DQHRCEBDVDCILEPSPQRCVNTQGGFECVNPYLDVDCVCEVDPDFRANCEYQC 360
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPNTQASCEPEGYLLDDGFI 438
DB 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPNTQASCEPEGYLLDDGFI 420
QY 439 CTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDCDGSKVDGSDSGEPPSPS 498
DB 421 CTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDCDGSKVDGSDSGEPPSPS 480
QY 499 PTPGSLTTPAVGLVHSG 516
DB 481 PTPGSLTTPAVGLVHSG 498
RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625

GENERAL INFORMATION:
; APPLICANT: Carlson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRE B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-312-870-3

Query Match 96.7%; Score 2820; DB 1; Length 497;
Best Local Similarity 99.8%; Pred. No. 5.2e-192;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 78
DB 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 60
QY 79 GGVGRRRLWIGLQPLPGCGDPRKRLGFRQWVTGNNTSYSRWALDLNGAPLCGLCV 138
DB 61 GGVGRRRLWIGLQPLPGCGDPRKRLGFRQWVTGNNTSYSRWALDLNGAPLCGLCV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPAGAAAVSITYGTPFA 198
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPAGAAAVSITYGTPFA 180
QY 199 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAGAWDCSVENGSGEACNA 258
DB 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAGAWDCSVENGSGEACNA 240
QY 259 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPSYSCMCTGYRLAA 318
DB 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPSYSCMCTGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPCPQRCVNTGGFECYCNVLDVGCVPVPDPQCFRANCEYQC 378
DB 301 DQHRCEVDVDCILEPSPCFQRCVNTGGFECYCNVLDVGCVPVPDPQCFRANCEYQC 360
QY 379 QPLNQTYSYLCVCAEGFAPIPHEPHRCQFCNQTACPADCDPNTQASCEPEGYILDGFI 438
DB 361 QPLNQTYSYLCVCAEGFAPIPHEPHRCQFCNQTACPADCDPNTQASCEPEGYILDGFI 420
QY 439 CTTDIBCEGNGFCGVCVCHNLPGTFECICGPDGSAVLRHIGTDCDSGVKVDGSGSGEPPPS 498
DB 421 CTTDIBCEGNGFCGVCVCHNLPGTFECICGPDGSAVLRHIGTDCDSGVKVDGSGSGEPPPS 480

QY 499 PTPGSTLTTPPAVGLVHS 515
DB 481 PTPGSTLTTPPAVGLVHS 497
RESULT 8
US-08-014-723-14
; Sequence 14, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-014-723-14

Query Match 95.0%; Score 2770; DB 1; Length 494;
Best Local Similarity 99.4%; Pred. No. 1.8e-188;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MLGVLVLGALALAGLPAPAEPOPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLM 60
DB 1 MLGVLVLGALALAGLPAPAEPOPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGNNTSYS 120
QY 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAP 240
QY 241 GAWDCSVENGSGEACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300

```
Db 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
QY 361 ECVEPVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 9
US-08-110-011A-14
; Sequence 14, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110.011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-14

Query Match 95.0%; Score 2770; DB 1; Length 494;
Best Local Similarity 99.4%; Pred. No. 1.8e-188;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVVLVGLALAGLGFAPAPQPGSQVCYHDCFALYPCGPFALNASQICDGLRHLM 60
Db 1 MGVVLVGLALAGLGFAPAPQPGSQVCYHDCFALYPCGPFALNASQICDGLRHLM 60
```

```
QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQPLPGCGDPRKLGPLRGFOWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQPLPGCGDPRKLGPLRGFOWVTGDNNTSYS 120
QY 121 RWARDLNGAPLCGPLCAVSAEAATVPSEPIWEEQOCEVKADGFLCEHFFPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCAVSAEAATVPSEPIWEEQOCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EFGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
Db 181 EFGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
QY 361 ECVEPVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 10
US-08-014-723-16
; Sequence 16, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014.723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-014-723-16

Query Match 94.9%; Score 2768; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MLGVLVLGALALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQGSYSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHYCNIDYLDVG 360
DB 301 DQGSYSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHYCNIDYLDVG 360
QY 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSGD 494

RESULT 11
US-08-110-011A-16
Sequence 16, Application US/08110011A
Patent No. 5354664

GENERAL INFORMATION:
APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Okuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110.011A
FILING DATE: 23-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5354664man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-073-0 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-110-011A-16

Query Match 94.9%; Score 2768; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MLGVLVLGALALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQGSYSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHYCNIDYLDVG 360
DB 301 DQGSYSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHYCNIDYLDVG 360
QY 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSGD 494

RESULT 12
US-08-307-444A-1
Sequence 1, Application US/08307444A
Patent No. 5516659

GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27

QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCEYGLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCEYGLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALVRH 475
Db 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALVRH 475

RESULT 14

US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-307-444A-2

Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1.5e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGFPAEPQGGSCQVEHDCFPALYFGPATFLNASQICDGLRGLHLM 60

Db 1 MLGVLVGLALAGLGFPAEPQGGSCQVEHDCFPALYFGPATFLNASQICDGLRGLHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCCDPRKRLGRLGFQWVTGNNNTSYS 120
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCCDPRKRLGRLGFQWVTGNNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCVASAAEAATVPSPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGPLCVASAAEAATVPSPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVQGHWAREAP 240
Db 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCEYGLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCEYGLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALVRH 475
Db 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALVRH 475

RESULT 15

US-08-587-389-2
; Sequence 2, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-587-389-2

```
Query Match          92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1.5e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLGVLVLGALALAGLGFAPAPQPGGQCVEHDCFALYPGPATFELNASQICDGLRGHLM 60
Db 1 MLGVLVLGALALAGLGFAPAPQPGGQCVEHDCFALYPGPATFELNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPPCGDPPKRLGRLRGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPPCGDPPKRLGRLRGFQWVTGDNNTSYS 120

Qy 121 RWRDLNLGAPLCGLPCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEFFHPATCRPLAV 180
Db 121 RWRDLNLGAPLCGLPCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEFFHPATCRPLAV 180

Qy 181 EPGAAAAVSTYGTPEAARGADFOALPVGSSAAVAPLGLQLMCTAPPCAVQGHWAREAP 240
Db 181 EPGAAAAVSTYGTPEAARGADFOALPVGSSAAVAPLGLQLMCTAPPCAVQGHWAREAP 240

Qy 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300

Qy 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILLEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILLEPSPQRCVNTQGGFECCHCYPNYDLVDG 360

Qy 361 ECVEVPDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
Db 361 ECVEVPDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420

Qy 421 TQASCEPGEYLLDDGFICTDIDECENGCGVCHNLPGTFECICGPDLSALVRH 475
Db 421 TQASCEPGEYLLDDGFICTDIDECENGCGVCHNLPGTFECICGPDLSALVRH 475
```

Search completed: May 8, 2002, 12:39:20
Job time: 211 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:19 ; Search time 42.86 Seconds
(without alignments)
891.783 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MLGVLVGLALAGLGFAP.....PSPTPGSTLTPPAVGLVHSG 516

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_1101.*

```
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2916	100.0	516	13 AAR22016	Truncated human th
2	2916	100.0	516	20 AAY09347	Human thrombomodul
3	2916	100.0	516	21 AAY83934	Human thrombomodul
4	2916	100.0	516	21 AAY69529	Human thrombomodul
5	2916	100.0	575	12 AAR11534	Human thrombomodul
6	2916	100.0	575	14 AAR41806	Thrombomodulin. H
7	2916	100.0	575	14 AAR43031	Human thrombomodul
8	2912	99.9	516	13 AAR22018	Human thrombomodul
9	2912	99.9	516	20 AAY09348	Human thrombomodul
10	2912	99.9	516	21 AAY83935	Human thrombomodul
11	2908	99.7	516	13 AAR22017	Human thrombomodul

12	2906	99.7	515	12	AAR10617	Soluble thrombomod
13	2906	99.7	575	9	AAP82070	Human thrombomodul
14	2904	99.6	575	14	AAR31572	Human thrombomodul
15	2899	99.4	575	13	AAR20639	Human urinary thro
16	2875	98.6	575	20	AAW73970	Human thrombomodul
17	2858	98.0	575	13	AAR22189	Sequence of thromb
18	2830	97.1	498	16	AAR84185	Human derived thro
19	2830	97.1	498	18	AAW01600	Thrombomodulin TME
20	2825	96.9	498	21	AAV67402	Novel sugar chain-
21	2824	96.8	497	19	AAW69520	rSTM protein SEQ I
22	2820	96.7	497	17	AAV94607	Human recombinant
23	2803	96.1	494	21	AAV67401	Novel sugar chain-
24	2803	96.1	500	21	AAV69530	Human thrombomodul
25	2770	95.0	494	16	AAV78726	Thrombomodulin wit
26	2768	94.9	494	13	AAR24400	Recombinant thromb
27	2753	94.4	494	16	AAR78727	Thrombomodulin wit
28	2748	94.2	486	12	AAR13877	Thrombin-binding s
29	2720	93.3	480	13	AAR22013	Truncated human th
30	2690	92.2	475	13	AAR22032	Mature thrombomodul
31	2684	92.0	476	16	AAR78725	Modified thrombomo
32	2684	92.0	476	16	AAR86376	Modified thrombomo
33	2657	91.1	535	9	AAV80641	Sequence of human
34	2619	89.8	462	15	AAR45336	Thrombomodulin ana
35	2615	89.7	462	15	AAR45337	Thrombomodulin ana
36	2615	89.7	462	15	AAR45338	Thrombomodulin ana
37	2598	89.1	462	15	AAR45339	Thrombomodulin ana
38	2596	89.0	462	15	AAR45350	Thrombomodulin ana
39	2591	88.9	462	15	AAR45343	Thrombomodulin ana
40	2572	88.2	462	15	AAR45349	Thrombomodulin ana
41	2571	88.2	462	15	AAR45353	Thrombomodulin ana
42	2566	88.0	462	15	AAR45347	Thrombomodulin ana
43	2565.5	88.0	461	15	AAR45335	Thrombomodulin ana
44	2561	87.8	462	15	AAR45346	Thrombomodulin ana

ALIGNMENTS

RESULT 1
AAR22016 *
ID AAR22016 standard; Protein; 516 AA.
XX
AC AAR22016;
XX
DT 03-JUL-1992 (first entry)
XX
DE Truncated human thrombomodulin encoded by plasmid pSV2TMD1.
XX
KW Thrombin binding site; blood clotting; TMD1 deleter.
XX
OS Homo sapiens.
XX
PN EP474273-A.
XX
PD 11-MAR-1992.
XX
PF 05-AUG-1991; 91EP-0202009.
XX
PR 03-AUG-1990; 90JP-0204978.
XX
PA (ASAH) ASahi KASEI KOGYO.
XX
Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
WPI; 1992-081820/11.
PT New polypeptide inhibits blood coagulation and platelet
aggregation - promotes thrombin catalysed protein C activation
for treating myocardial infarction, thrombosis, embolism, etc.
XX
PS Example 1; Page 17; 112pp; English.
XX

CC Plasmid M13mp19TMJ3 (see AAR22013) was used as template for
CC site-directed mutagenesis. A 177bp fragment was deleted using the
CC "TMd1" oligonucleotide to give plasmid M13TMd1 which encodes
CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid
CC M13TMd1 was completely digested with HindIII and BamHI and a TMd1
CC fragment of ca. 1700bp was isolated. The fragment was ligated to
CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant
CC plasmid pSV2TMd1. See AAR22014-R22022 and AAQ25072.
XX
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 13; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFAPAPAPQPGSQCVHDFCFALYPGPATFNASQICDGLRGHLM 60
DB 1 mlgvlvlgallaglgfpapapqpgsqcvhdfcfalypgpattfnasqicdglrghlm 60

QY 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFQWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppgcgdpkrlgpgfqwvtgdnntsys 120

QY 121 RWARLDNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFATCRPLAV 180
DB 121 rwarldngaplcgplcvavsaeatvpsepiweeqcevkadgflcefhfpatcrplav 180

QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaasvitygtptfaargadfoalpvgssaaavaplgqlmctappagvqghwareap 240

QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgaprcqcpagaaalqadgrsctasatqscndlcehfcvnp 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHECHYPNVDLVDG 360
DB 301 dqpgsyscmcetgyrllaadqhrcevdvdcilepspcqrcvntqggfdecheypnvdlvdg 360

QY 361 ECVEPVDPFCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
DB 361 ecvepvdpfcfranceyqcplnqtsylcvcaegfapiphephrcqmfncotacpadcdpn 420

QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDC 480
DB 421 tqascepegyllddgfictdidecengfcgvcvchnlpgtfecicgpdssalvrhigtgc 480

QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAGVLVHSG 516
DB 481 dsgkvdgds gsgseppptpgstltpppagvlvhs g 516

RESULT 2
AA09347
ID AA09347 standard; Protein; 516 AA.
XX
XX
AC AA09347;
DT 08-JUL-1999 (first entry)
XX
XX Human thrombomodulin SEQ ID NO:1.
DE
XX Human; thrombomodulin; aqueous parenteral solution; storage;
KW distribution; acute coronary syndrome; thrombosis; embolism;
KW diabetes.
XX
XX Homo sapiens.
OS
XX
PN WO9918994-A1.
XX
PD 22-APR-1999.
XX

PF 13-OCT-1998; 98WO-JP04609.
XX
PR 11-NOV-1997; 97JP-0308523.
PR 15-OCT-1997; 97JP-0281659.
XX
PA (ASAH) ASAH KASEI KOGYO KK.
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
XX
XX WPT; 1999-277444/23.
DR N-PSDB; AAX55879.
XX
PT Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,
PT thrombosis, embolism, and diabetes
XX
PS Claim 6; Page 87-89; 97pp; Japanese.
XX
CC The present invention describes a method for maintaining the quality of
CC an aqueous parenteral solution of thrombomodulin comprising buffer and
CC surfactant aseptically filled in a case or syringe. Maintaining the
CC quality of an aqueous, parenteral thrombomodulin solution is
CC characterised in that the solution: (a) comprises soluble thrombomodulin,
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
CC filled into: (i) a case or (ii) a syringe without any empty space; and
CC (c) is kept in liquid form in storage and distribution and not frozen or
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome
CC (such as myocardial infarction, unstable angina and coronary artery
CC vessel) thrombosis (e.g. cerebral, vascular and peripheral blood
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
CC lupus erythematosus or Barrett's syndrome), multiple organ failure,
CC disseminated intravascular coagulation, transient ischaemia, diabetes,
CC liver veno-occlusive diseases and deep vein thrombosis. The composition
CC is stable for a long period of time and can be stored and distributed in
CC ready to use form avoiding the problems of dissolution and accuracy when
CC preparing on demand. The present sequence represents human
CC thrombomodulin.
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFAPAPAPQPGSQCVHDFCFALYPGPATFNASQICDGLRGHLM 60
DB 1 mlgvlvlgallaglgfpapapqpgsqcvhdfcfalypgpattfnasqicdglrghlm 60

QY 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFQWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppgcgdpkrlgpgfqwvtgdnntsys 120

QY 121 RWARLDNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFATCRPLAV 180
DB 121 rwarldngaplcgplcvavsaeatvpsepiweeqcevkadgflcefhfpatcrplav 180

QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaasvitygtptfaargadfoalpvgssaaavaplgqlmctappagvqghwareap 240

QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgaprcqcpagaaalqadgrsctasatqscndlcehfcvnp 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHECHYPNVDLVDG 360
DB 301 dqpgsyscmcetgyrllaadqhrcevdvdcilepspcqrcvntqggfdecheypnvdlvdg 360

QY 361 ECVEPVDPFCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
DB 361 ecvepvdpfcfranceyqcplnqtsylcvcaegfapiphephrcqmfncotacpadcdpn 420

QY 421 TQASCECEGYILLDDGFICTDIDECENGRCFCGVCVCHNLPGTFCICGPDLSALVRHIGTDC 480
DB 421 tqascecegyillddgfictdidecengrcfcgvcvchnlpgtfcicgpdlsalvrhigtcd 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
DB 481 dsqkvdgsgsgsgseppspptpgstltppavglvhsq 516

RESULT 3
AAV83934
ID AAY83934 standard; Protein; 516 AA.
XX AC AAY83934;
XX DT 28-JUL-2000 (first entry)
XX DE Human thrombomodulin TMD protein.
XX KW Human; thrombomodulin; vasculitis; protein C; thrombin.
XX OS Homo sapiens.
XX PN JP2000053582-A.
XX PD 22-FEB-2000-
XX PF 06-AUG-1998; 98JP-0222688.
XX PR 06-AUG-1998; 98JP-0222688.
XX PA (ASAH) ASAH KASEI KOGYO KK.
XX DR WPI: 2000-353249/31.
XX DR N-PSDB; AAA10027.
XX PT Treating agent for vasculitis contains peptide which promotes
XX PT activation of protein C by thrombin -
XX PS Claim 4; Page 10-12; 18pp; Japanese.
XX CC This sequence represents a human thrombomodulin protein. The invention
XX CC relates to a treating agent for vasculitis containing a peptide which
XX CC promotes activation of protein C by thrombin.
XX SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGPAPAEPOPGSGQCVHEHDFALYPGPATFLNASQICDGLRGLHM 60
DB 1 mlgvvlvlgalaalaglgfpapaeppgsgqcvhehdfalypgpaflnasqicdglrghlm 60
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120
DB 61 tvrssvaadvlsllngdggvgrrrlwiglppcgdprrklrgpfqwtgdnntsys 120
QY 121 RWARDLNGAPLCGPLCVAVSAAATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 rwardlngaplcgplcvavsaaatvpsepiweeqcevkadgflcefhfpatcrplav 180
QY 181 EPGAAAASVTYCTPBAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGQGHWAREAP 240
DB 181 epgaaaasvtyctpbaargadfoalpvgssaaavaplglqlmctappvavqghwareap 240
QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgaprcqcpagaaqadgrsctasatqscndlcehfcvnp 300

QY 301 DQPGSYSCMCTGYRLAADHRCEDVDDCILBSPCPQRCVNTQGGFECGHCYPNVLDVDS 360
DB 301 dqpgsyscmctgyrllaadhrcedvddcilbpspcqrcvntqggfegchcypnyldvdg 360
QY 361 ECVEPYDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQFMCNQTACPADCDPN 420
DB 361 ecvepydpcfranceyqcqplnqtsylcvcaegfapiphephrcqfmcnqtacpadcdpn 420
QY 421 TQASCECEGYILLDDGFICTDIDECENGRCFCGVCVCHNLPGTFCICGPDLSALVRHIGTDC 480
DB 421 tqascecegyillddgfictdidecengrcfcgvcvchnlpgtfcicgpdlsalvrhigtcd 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516
DB 481 dsqkvdgsgsgsgseppspptpgstltppavglvhsq 516

RESULT 4
AAV69529
ID AAY69529 standard; Protein; 516 AA.
XX AC AAY69529;
XX DT 10-APR-2000 (first entry)
XX DE Human thrombomodulin variant, SEQ ID NO:3.
XX KW Thrombomodulin; TM; soluble; affinity purification; cation exchange;
XX KW anticoagulant; thrombolytic.
XX OS Homo sapiens.
XX PN JP11341990-A.
XX PD 14-DEC-1999.
XX PF 23-MAR-1999; 99JP-0077518.
XX PR 30-MAR-1998; 98JP-0084389.
XX PA (ASAH) ASAH KASEI KOGYO KK.
XX DR WPI: 2000-101696/09.
XX DR N-PSDB; AAZ55965.
XX PT Preparation of highly pure soluble thrombomodulin - used as an
XX PT antiblood coagulation agent and a thrombolytic agent
XX PS Claim 9; Page 30-32; 36pp; Japanese.
XX CC The invention relates to a novel method for the preparation of highly
XX CC pure soluble thrombomodulin (TM) containing substantially no serum-
XX CC derived or antibody-derived substance. The method comprises isolating
XX CC soluble TM from unpurified serum via affinity purification using an
XX CC anti-TM antibody. The soluble TM is then treated with a cation
XX CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to
XX CC 4, and the fraction containing the soluble TM is isolated. The soluble
XX CC TM can be used as a blood anticoagulation agent and a thrombolytic agent.
XX CC This sequence represents a human thrombomodulin variant, designated
XX CC SEQ ID NO:3.
XX SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGPAPAEPOPGSGQCVHEHDFALYPGPATFLNASQICDGLRGLHM 60
DB 1 mlgvvlvlgalaalaglgfpapaeppgsgqcvhehdfalypgpaflnasqicdglrghlm 60
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120

Db 61 tvrsvaadvislllmgdgvgrrrlwlgqlppgcgpkrlgplrgfqwtgdnntsys 120
Qy 121 RWARLDLNGALGCLGCLVAVSAATVPSEPIWEEQQCEVKADGFLCEHFHPCRCPLAV 180
Db 121 rwarldlngaplcgclvavsaatvpsepilweeqqcevkdadgflcehfhpacrcplav 180
Qy 181 EPGAAAAVSTYGTFFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 epgaavaavstytgtpfaargadqalpvgssaaavaplgqlmctappagvqghwareap 240
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcehacnaipgprcqcpgaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360
Db 301 dqpgsyscmetgyrllaadqhrcevdvdcillepspcprcvntqggfchcypnydlvdg 360
Qy 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvpdpcfranceyqcqplnqtsylcvcaegfapiphehrcqmfncqtacpadcnp 420
Qy 421 TQASCECPGYLLDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecepgyllddgfictdidecengfgcgvchnlpgtfecicgpdlsalvrhigtcd 480
Qy 481 DSGKVDGDSGSGEPPPTGSLTTPPAVGLVHSG 516
Db 481 dsgkvdgdsdgsgeppptgstlttppavglvhs 516

RESULT 5

AAR11534
ID AAR11534 standard; Protein; 575 AA.

XX AAR11534;

XX 17-JUN-1991 (first entry)

XX Human thrombomodulin type II polypeptide.

XX Thrombomodulin; chondroitin; protein C; thrombin; thrombolysis;
XX chondroitin sulphate; anticoagulant; myocardial infarction.

XX Key Location/Qualifiers

FT - Peptide 1..18

FT - Peptide /label= signal peptide

FT - Peptide 19..365

FT - Peptide /label= opt. N-terminal sequence

FT - Peptide 366..480

FT - Peptide /label= human thrombomodulin active fragment

FT - Peptide 481..516

FT - Peptide /label= opt. C-terminal sequence

XX WO9104276-A.

XX 04-APR-1991.

XX 25-SEP-1990; 90WO-JP01234.

XX 25-SEP-1989; 89JP-0246270.

XX (ASAH) ASAH KASEI KOGYO.

XX Yamamoto S, Gomi K, Ogawa K;

XX WPI; 1991-117478/16.

XX New human thrombomodulin polypeptide contg. chondroitin - has
XX anticoagulant platelet agglutination and thrombolytic activity

XX Disclosure; fig 1; 80pp; Japanese.

XX This human thrombomodulin polypeptide comprises a 115 residue active
CC fragment with opt. flanking sequences (N-terminal= 347 amino acids;
CC C-terminal= 36 amino acids). Additionally it has an N-terminal signal
CC sequence and an attached sugar chain, contg. chondroitin or chondroitin
CC sulphate. It promotes the activation of protein C by thrombin and has
CC good anticoagulant, platelet aggregation inhibition and thrombolytic
CC activities. It can therefore be used for treating cardiovascular
CC diseases, eg arteriosclerosis or myocardial infarction.

XX Sequence 575 AA;

Query Match 100.0%; Score 2916; DB 12; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLGVLVLGALALAGLGFAPAEPPQPGSQCVHEHDFALYPGATFLNASQICDGLRHLM 60
Db 1 mlgvllvlgalalaglgfpapaeppqpgsqcvhehdfalypgpatflnasqicdglrhl 60

Qy 61 TVRSSVAADVISLLLNGDGGVGRRLWLGQLPPCGDPKRLGPLRGFQWVTGDNNTSYS 120
Db 61 tvrsvaadvislllmgdgvgrrrlwlgqlppcgdpkrlgplrgfqwtgdnntsys 120

Qy 121 RWARLDLNGALGCLGCLVAVSAATVPSEPIWEEQQCEVKADGFLCEHFHPCRCPLAV 180
Db 121 rwarldlngaplcgclvavsaatvpsepilweeqqcevkdadgflcehfhpacrcplav 180

Qy 181 EPGAAAAVSTYGTFFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 epgaavaavstytgtpfaargadqalpvgssaaavaplgqlmctappagvqghwareap 240

Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcehacnaipgprcqcpgaalqadgrsctasatqscndlcehfcvnp 300

Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360
Db 301 dqpgsyscmetgyrllaadqhrcevdvdcillepspcprcvntqggfchcypnydlvdg 360

Qy 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvpdpcfranceyqcqplnqtsylcvcaegfapiphehrcqmfncqtacpadcnp 420

Qy 421 TQASCECPGYLLDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecepgyllddgfictdidecengfgcgvchnlpgtfecicgpdlsalvrhigtcd 480

Qy 481 DSGKVDGDSGSGEPPPTGSLTTPPAVGLVHSG 516

Db 481 dsgkvdgdsdgsgeppptgstlttppavglvhs 516

RESULT 6

AAR41806
ID AAR41806 standard; peptide; 575 AA.

XX AAR41806;

XX 30-MAR-1994 (first entry)

XX Thrombomodulin.

XX Transformation; fungus; blood coagulation; prevention; platelet;
KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;
KW site-directed mutagenesis; promotion; protein C; activation; thrombin.

XX Homo sapiens.

XX JP05213998-A.

XX 24-AUG-1993.

```
XX 02-AUG-1991; 91JP-0282369.
XX
XX 03-AUG-1990; 90JP-0204978.
PR 30-JUL-1991; 91JP-0189984.
XX
XX (ASAH ) ASahi CHEM IND CO LTD.
PA
XX WPI; 1993-299652/38.
XX
XX Novel polypeptide obtd. by culturing transformed fungus - having
PT blood coagulation preventing, platelet aggregation preventing and
PT thrombolytic activities
XX
XX Disclosure; Fig 55; 65pp; Japanese.
XX
XX Novel polypeptides, obtd. by culturing transformed fungus, have
CC blood coagulation preventing, platelet aggregation preventing
CC and thrombolytic activities
CC
CC In an example, plasmid M13mpl19TMD3 (constructed from pSV2TMD2
CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected
CC to site directed mutagenesis to prepare pSV2TMD7. plasmid pSV2TMD7
CC was transfected to COS-1 cells. The activity of promoting protein C
CC activation by thrombin of the peptide produced by the transformed
CC COS-1 cell was measured. The amt. of the peptide was determined.
XX
XX Sequence 575 AA;
SQ
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSQGVHDCFALYPGPAPFLNASQICDGLRGHLM 60
DB 1 mlglvlgallalaglgfpapepqpqgscvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPGCCGDKPKRLGPRGFWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppgccgdkprlgfgrfwvtdnntsys 120
QY 121 RWARLDLNGAPLCPCLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHPCPLAV 180
DB 121 rwarldlngaplcgplcvavsaeeatvpsepieweeqcevkdgflcehfpcplav 180
QY 181 EPGA AAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 epga aavstytgtptfaargadfoalpvgsaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVPNP 300
DB 241 gawdcsvengcgcehacnaipgaprcqcpagaaqlqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
DB 301 dqpgsyscmctgyrlaadqhrcevdvdcilep spqrcvntqggftechchypnydlvdg 360
QY 361 ECVEPVDPFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMCQNTACADCDPN 420
DB 361 ecvepvdpfranceyqcqplnqtsylcvcaegfapiphephrcmqntacpdcadcpn 420
QY 421 TQASCEPEGYILDGFGTCTDIDECENGFGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 tqascepegyildgfgtctdidcengfgsgvchnlpgtfecicgpdlsalvrhigtdc 480
QY 481 DSGKVDGSDGSGPSPPTPGSLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgdsgspptpgsltltpavglvhs 516
RESULT 7
AAR43031
ID AAR43031 standard; protein; 575 AA.
```

```
XX AAR43031;
AC
XX
XX 16-MAY-1994 (first entry)
XX
XX Human thrombomodulin.
XX
XX Anticoagulant; platelet aggregation inhibitor;
KW protein C; activation; thrombin; thrombomodulin;
KW coagulation disorder; thrombosis; myocardial infarction;
KW embolism; telangiectasis; arteriosclerosis obliterans;
KW disseminated intravascular coagulation; DIC; angina pectoris;
KW gestosis; transient ischaemic attack.
XX
XX Homo sapiens.
XX
XX WO9322447-A.
XX
XX 11-NOV-1993.
XX
XX 30-APR-1993; 93WO-JP00578.
XX
XX 01-MAY-1992; 92JP-0112903.
XX
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX
XX Kondo S, Toma K, Zushi M;
XX
XX WPI; 1993-368806/46.
XX
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX
XX Disclosure; Fig 1; 84pp; Japanese.
XX
XX New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX
XX Sequence 575 AA;
SQ
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSQGVHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlglvlgallalaglgfpapepqpqgscvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPGCCGDKPKRLGPRGFWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppgccgdkprlgfgrfwvtdnntsys 120
QY 121 RWARLDLNGAPLCPCLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHPCPLAV 180
DB 121 rwarldlngaplcgplcvavsaeeatvpsepieweeqcevkdgflcehfpcplav 180
QY 181 EPGA AAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 epga aavstytgtptfaargadfoalpvgsaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVPNP 300
DB 241 gawdcsvengcgcehacnaipgaprcqcpagaaqlqadgrsctasatqscndlcehfcvnp 300
```

QY	301	DQPSYS	CMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECYENYDLVDG	360
Db	301	dqpsys	cmctgyrllaadqhrcevdadcilepspcqrcvntqggfchecyenydlvdg	360
QY	361	ECVEPVD	PCFRANCEYQCQPLNQTSYLCVCAEGFAPITPHEPHRCQMFNCQTACPADCDPN	420
Db	361	ecvepvd	pcfranceyqcqplnqtsylcvcaegfapitphephrcqmfncqtacpdcqpn	420
QY	421	TOASCE	PEGVILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDPSALVRHIGTDC	480
Db	421	tgasce	pegvillddgfictdidecengfcsgvchnlpgtfecicgpdpsalvrhigtgc	480
QY	481	DSGKVDG	SGSGEPPTPGSTLTTPPAVGLVHSG	516
Db	481	dsgkvdg	sgsgseppspcpgscittpavglvns	516
	RESULT	8		
	AAAR22018			
ID	AAAR22018	standard;	Protein; 516 AA.	
XX	AAAR22018;			
XX				
DT	03-JUL-1992	(first entry)		
XX				
DE	Human thrombomodulin (1-516) with Asp367 substituted by Glu.			
XX				
KW	Mutant; thrombin binding site; blood clotting; Tmm2 mutator.			
XX				
OS	Homo sapiens.			
XX				
PN	EP474273-A.			
XX				
PD	11-MAR-1992.			
XX				
PF	05-AUG-1991; 91EP-0202009.			
XX				
PR	03-AUG-1990; 90JP-0204978.			
XX				
PA	(ASAH) ASahi Kasei Kogyo.			
XX				
PI	Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;			
XX				
DR	WPI; 1992-081820/11.			
XX				
PT	New polypeptide inhibits blood coagulation and platelet			
PT	aggregation - promotes thrombin catalysed protein C activation			
PT	for treating myocardial infarction, thrombosis, embolism, etc.			
XX				
PS	Example 1; Page 18; 112pp; English.			
XX				
CC	Plasmid M13PMD1 (see AAR22016) encoding the first 516 N-terminal			
CC	amino acids of human thrombomodulin was used as a template for			
CC	site-directed mutagenesis using the "Tmm2 mutator" to produce			
CC	plasmid M13PMM2. In the mutant thrombomodulin encoded by the			
CC	plasmid, the wild-type Asp residue at position 367 is substituted			
CC	by an Glu residue. The activity of this mutant, truncated			
CC	thrombomodulin (i.e. D123Glu) in the activation of protein C was			
CC	compared to that of similarly truncated thrombomodulin with Asp			
CC	at position 367 (i.e. D123Asp). The activity of D123Glu was as high			
CC	as 1.2 times that of D123Asp. See AAR22013-R22022 and AAQ25074.			
XX				
SQ	Sequence	516 AA;		

QY	61	TVRSSVAADVISLLNGDGGVRRRLTIGLQLPGCCDPKRIPLRGFQWVTGDNNTSYS	120
Db	61	tvrssvaadvvisllngdggvrrrlwiglqlppgcgdkrllgplrgfwvtgdnntsys	120
QY	121	RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAV	180
Db	121	rwarldlngapl.cplplcvavsaaeatvpsepiweeqcevkadgflcehfhpacrlav	180
QY	181	EPGAAAAVSYTGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Db	181	epgaaaaavsytygtptfaargadfgalpvgsaaavaplgllmctappgaavgghwareap	240
QY	241	GAWDCSVENGCEHACNAIFCAPRCQCPAGAAIQADGRSCTASATOSCNDLCEHFCVPNP	300
Db	241	gawdcsvengcehacnaifgaprcqcpagaaiaqadgrsctasatqscndlcehfvcvnp	300
QY	301	DQPGSYSQCMCTGYRLAADQHRCEVDVDCILEPSPCPCRCVNTQGGFECGCHYPNYDLVDG	360
Db	301	dqpgsysqcmctgyrllaadqhrcevdvdcilepspcprcvntqggfecgchypnydlvdg	360
QY	361	ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQFCNQTCAPADCDPN	420
Db	361	ecvepvcfranceyqcqplnqtsylcvcaegfapiphephrcqfcmfcnqtacpdcdpn	420
QY	421	TQACECEPEGYILDGFGICTDIDBCENGGFCGVCNHLPGTFECICGPPDSALVRHIGTDC	480
Db	421	tqacecepegylldgfgictdidbceנגgfcsgvchnlpgtfeclcgppdsalvrhigtcd	480
QY	481	DSGKVDGSDSGSRRPPTPGSLTPPAVGLVHSG	516
Db	481	dsgkvdgsgdsgeppspptpgsltpavglvhs	516
RESULT	9		
AAAY09348			
ID	AAAY09348	standard; Protein; 516 AA.	
AC	AAAY09348;		
XX			
DT	08-JUL-1999	(first entry)	
XX			
DE	Human thrombomodulin	SEQ ID NO:2.	
XX			
KW	Human; thrombomodulin; aqueous parenteral solution; storage;		
KW	distribution; acute coronary syndrome; thrombosis; embolism;		
XX	diabetes.		
OS	Homo sapiens.		
XX			
PN	W09918994-A1.		
XX			
PD	22-APR-1999.		
XX			
PF	13-OCT-1998; 98WO-JP04609.		
XX			
PR	11-NOV-1997; 97JP-0308523.		
PR	15-OCT-1997; 97JP-0281659.		
XX			
PA-	(ASAH) ASAHI KASEI KOGYO KK.		
XX			
PI	Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;		
XX			
DR	WPI: 1999-277444/23.		
DR	N-PSDB; AAX55880.		
XX			
PT	Stable aqueous parenteral thrombomodulin solution - comprising		
PT	buffer and surfactant, useful for treating acute coronary syndrome,		
PT	thrombosis, embolism, and diabetes		
XX			
PS	Claim 7; Page 90-92; 97pp; Japanese.		
XX			
CC	The present invention describes a method for maintaining the quality of		

Query Match 99.7%; Score 2906; DB 12; Length 515;
Best Local Similarity 99.8%; Pred. No. 8.3e-162;
Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLGVLVIGALALAGLGFAPAPAEPPQGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 mlgvlvigalalaglgfpapaeppqgscqvehdcfalyppgpatflnasqicdglrghlm 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPRGFQWVTGDNNTSYS 120
Db 61 tvrssvaadvvisllngdggvrrrlwiglqlppgcgdpkrlgprgfwvtdgdnntsys 120
Qy 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 180
Db 121 rwarldlngaplcgplcvavsaeeatvpsepiweeqqcevkadgflcehfhpaterplav 180
Qy 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 epgaaaasvitygtptfaargadfoalpvgssaaavaplgqlmctappgavqghwareap 240
Qy 241 GAWDCSVENGCGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcgcehacnaipgprcqpagaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
Db 301 dqpgsyscmctgyrllaadhrcedvddcilepspcrcvntqggfchecypnydlvdg 360
Qy 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvdpcfranceyqcplnqtsylcvcaeafapiphephrcqmfncqtacpadcdpn 420
Qy 421 TOASCCEPEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
Db 421 tqascecepegyllddgfictdidengfcgvcnhlpgtfecicgpdalsalvrhigtcd 480
Qy 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHVS 515
Db 481 dsgkvdgsgsgseppptpgstlttppavglvhs 515

RESULT 13
AAP82070
ID AAP82070 standard; protein; 575 AA.
XX
AC AAP82070;
DT 19-OCT-1990 (first entry)
DE Human thrombomodulin encoded by plasmid p2.1.
DE
KW thrombomodulin activity; Protein C; anticoagulant;
KW epidermal Growth Factor (EGF) domains.
XX
OS synthetic.
XX
XX W08809811-A.
XX
PD 15-DEC-1988.
XX
PF 09-JUN-1988; 88WO-DK00089.
XX
PR 12-JUN-1987; 87DK-0002990.
XX
PA (NOVO) NOVO INDUSTRI A/S.
XX
XX Nexo BA, Esper B;
XX
XX WPI; 1988-368626/51.
DR N-PSDB; AAP82026.
XX

PT Recombinant protein having thrombomodulin activity -
PT used in the therapeutic control of coagulation and the treatment
PT and prevent of thrombotic episodes
XX
PS Disclosure; ; p; English.
XX
CC Protein is encoded by plasmid p2.1 derived from human cell
CC line A549 known to express about 10000 molecules of thrombomodulin
CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a
CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
CC potentiate a patient's anticoagulant capacity.
CC See also AAN82027 and AAN82037.
XX
SQ Sequence 575 AA;

Query Match 99.7%; Score 2906; DB 9; Length 575;
Best Local Similarity 99.6%; Pred. No. 9.2e-162;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLGVLVIGALALAGLGFAPAPAEPPQGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 mlgvlvigalalaglgfpapaeppqgscqvehdcfalyppgpatflnasqicdglrghlm 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPRGFQWVTGDNNTSYS 120
Db 61 tvrssvaadvvisllngdggvrrrlwiglqlppgcgdpkrlgprgfwvtdgdnntsys 120
Qy 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 180
Db 121 rwarldlngaplcgplcvavsaeeatvpsepiweeqqcevkadaficehfhpaterplav 180
Qy 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 epgaaaasvitygtptfaargadfoalpvgssaaavaplgqlmctappgavqghwareap 240
Qy 241 GAWDCSVENGCGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcgcehacnaipgprcqpagaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
Db 301 dqpgsyscmctgyrllaadhrcedvddcilepspcrcvntqggfchecypnydlvdg 360
Qy 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvdpcfranceyqcplnqtsylcvcaeafapiphephrcqmfncqtacpadcdpn 420
Qy 421 TOASCCEPEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
Db 421 tqascecepegyllddgfictdidengfcgvcnhlpgtfecicgpdalsalvrhigtcd 480
Qy 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHVS 516
Db 481 dsgkvdgsgsgseppptpgstlttppavglvhs 516

RESULT 14
AAP821572
ID AAP821572 standard; Protein; 575 AA.
XX
AC AAP821572;
XX
DT 27-MAY-1993 (first entry)
DE Human thrombomodulin.
DE
KW Mutagenesis; site specific; nucleic acid constructs;
KW restriction site; introduction; removal.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

```
FT Peptide 1..20
FT Domain /note= "signal peptide"
FT 245..280
FT /note= "EGF-1"
FT 288..323
FT /note= "EGF-2"
FT 329..361
FT /note= "EGF-3"
FT 368..404
FT /note= "EGF-4"
FT 408..439
FT /note= "EGF-5"
FT 445..480
FT /note= "EGF-6"
FT 481..515
FT /note= "O-linked glycosylation domain"
FT Region 516..538
FT /note= "stop transfer sequence"
FT Domain 539..575
FT /note= "cytoplasmic domain"
XX
XX WO9301282-A.
XX
XX 21-JAN-1993.
XX
XX 01-JUL-1992; 92WO-US05573.
XX
XX 01-JUL-1991; 91US-0724237.
XX
XX (BERL-) BERLEX LAB INC.
XX
XX Andrews WH, Morser MJ, Vilander LR;
XX
XX WPI; 1993-045488/05.
XX
XX Site-specific mutagenesis of nucleic acid constructs - using an
XX oligo:nucleotide which changes a nucleotide and introduces or
XX removes a restriction site
XX
XX Example; Fig 3; 87pp; English.
XX
XX The sequence is that of the native human thrombomodulin showing the
XX six EGF-like domains.
XX
XX Sequence 575 AA;

Query Match 99.6%; Score 2904; DB 14; Length 575;
Best Local Similarity 99.6%; Pred. No. 1.2e-161;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLGVLVLGALALAGLGFAPAEPQGSGQCVHEHDFALYGPATFLNASQICDGLRGHLM 60
Db 1 mlgvvlvgalalaglgfpapepgpgsgqcvhehdcslalypapatflnasqicdglrghlm 60
Qy 61 TVRSVAADVLSLLNGDGGVRRRLWIGLQLPPCGDPRKRLGPRGFQWVTGDNNTSYS 120
Db 61 tvrsvaadvlsllngdggvrrrlwlgllppcgdpkrllgprlgfqrwvtgdnntsys 120
Qy 121 RWARLDLNGPLCGPLCVANSAATVPSEPIWEEQQCEYKADGFLCEFFHFPATCRPLAV 180
Db 121 rwarldngplcplcvansaavtvpsepiweeqqcevkadgflcefhfpatcrplav 180
Qy 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVOGHWAREAP 240
Db 181 epgaaaaavstytptfaargadfqalpvgsaaavaplgllmctappgavghwareap 240
Qy 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300
Db 241 gawdcsvenggehacnaipgaprcqcpagaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCQVNTGGFECHCYPNYDLVDG 360
Db 301 dpgsyscmctgyrllaadqhrcevdvdcilep spqrcqvntggfecchcypnydlvdg 360
```

```
Db 301 dpgsyscmctgyrllaadqhrcevdvdcilep spqrcqvntggfecchcypnydlvdg 360
Qy 361 ECVEPVPDFCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
Db 361 ecvepvpdfcfranceyqcqlnqtsylcvcaegfapiphephrcqmfcnqtacpadcdpn 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGCGFCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecenggcfcgvchnlpgtfecicgpdlsalarhigtcd 480
Qy 481 DSGKVDGDSGSGEPSPPTFGSILTPPAVGLVHSG 516
Db 481 dsgkvdgds gsgseppspptfgsiltp pavglvhs g 516

RESULT 15
AAR20639
ID AAR20639 standard; Protein; 575 AA.
XX
XX AC AAR20639;
XX
XX 01-MAY-1992 (first entry)
XX
XX Human urinary thrombomodulin.
XX
XX Blood clotting; intravascular coagulation.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..18
XX /label= signal
XX Protein 19..575
XX /label= thrombomodulin
XX
XX WO9200325-A.
XX
XX 09-JAN-1992.
XX
XX 27-JUN-1991; 91WO-JP00873.
XX
XX 27-JUN-1990; 90JP-0168766.
XX
XX (MOCH ) MOCHIDA PHARM KK.
XX
XX Nii A, Morishita H, Uemura A, Mochida E;
XX
XX WPI; 1992-041517/05.
XX
XX N-PSDB; AAQ20810.
XX
XX New modified recombinant human urinary thrombomodulin - has
XX thrombin binding, anticoagulant and thrombolytic activity and is
XX for treating blood coagulation disorders
XX
XX Claim 1; Fig 3; 107pp; Japanese.
XX
XX A 2.5kb cDNA fragment was isolated by screening a human placental
XX cell library with a probe coding for the N-terminal of human
XX thrombomodulin. This full-length sequence was modified to produce
XX two alternative, truncated forms of the coding sequence encoding
XX only up to amino acid 456 of the mature polypeptide; the Ala
XX residue at position 455 is substituted by Val in one of the two
XX truncated forms.
XX
XX Sequence 575 AA;

Query Match 99.4%; Score 2899; DB 13; Length 575;
Best Local Similarity 99.4%; Pred. No. 2.3e-161;
Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLGVLVLGALALAGLGFAPAEPQGSGQCVHEHDFALYGPATFLNASQICDGLRGHLM 60
Db 1 mlgvvlvgalalaglgfpapepgpgsgqcvhehdcslalypapatflnasqicdglrghlm 60
```

Db	1	m1gvlvlgalalaglgfpapaeppqpgsgqvehdcfalypgpatfinaasqicdglrgqlm	60
Qy	61	TVRSSVAADVISILLNGDGVGRRRLWIGLQLPPGCGDKRKLGLRGFOWVTGDNNTSYS	120
Db	61	TVRSSVADVISILLNGDGVGRRRLWIGLQLPPGCGDKRKLGLRGFOWVTGDNNTSYS	120
Qy	121	RWARLDLNGAPLCGPILCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAY	180
Db	121	RWARLDLNGAPLCGPILCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAY	180
Qy	181	EPGAAAAYSIYTGTFPARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240
Db	181	EPGAAAAYSIYTGTFPARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240
Qy	241	GAWDCSVENGCEHACNAIPGAPRCOCPAGALQADGRSCTASATOSCNLCEHFVCPNP	300
Db	241	GAWDCSVENGCEHACNAIPGAPRCOCPAGALQADGRSCTASATOSCNLCEHFVCPNP	300
Qy	301	DQPGSYSCMCTGYRLAABQHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
Db	301	DQPGSYSCMCTGYRLAABQHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
Qy	361	ECVEPVDFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN	420
Db	361	ECVEPVDFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN	420
Qy	421	TQASCECPGYILLDDGFICTDDECENGGFCSCGVCHNLPGTFECICGPDSSALVRHIGTDC	480
Db	421	TQASCECPGYILLDDGFICTDDECENGGFCSCGVCHNLPGTFECICGPDSSALVRHIGTDC	480
Qy	481	DSGKVDGDSGSGEPPPTPGSTLTPPAVGLVHSG	516
Db	481	DSGKVDGDSGSGEPPPTPGSTLTPPAVGLVHSG	516

Result No.	Query	Score	Match	Length	DB	ID	Description	
1	2916	100.0	575	1	TRBM_HUMAN	P07204	homo sapien	
2	1903	65.3	577	1	TRBM_MOUSE	P15306	mus musculus	
3	1095.5	37.6	356	1	TRBM_BOVIN	P06579	bos taurus	
4	374	12.8	2907	1	FBN2_MOUSE	Q61555	mus musculus	
5	368	12.6	2871	1	FBN1_BOVIN	P98133	bos taurus	
6	367.5	12.6	1184	1	FBF12_HUMAN	P98095	homo sapien	
7	366	12.6	2911	1	FBN1_PIG	Q9tv36	sus scrofa	
8	366	12.6	2911	1	FBN2_HUMAN	P35556	homo sapien	
9	364	12.5	2871	1	FBN1_HUMAN	P35555	homo sapien	
10	362	12.4	2871	1	FBN1_MOUSE	Q61554	mus musculus	
11	357	12.2	1221	1	FBF12_MOUSE	P37889	mus musculus	
12	356	12.2	956	1	FBF12_HUMAN	O00339	homo sapien	
13	332.5	11.4	1964	1	NTC4_MOUSE	P31695	mus musculus	
14	330.5	11.3	1712	1	TGFB_RAT	O00918	rattus norv	
15	326.5	11.2	956	1	FBN2_MOUSE	O08746	mus musculus	
16	324.5	11.1	712	1	FBF1_CAEEL	O77469	caenorhabdi	
17	322.5	11.1	1394	1	TGFB_HUMAN	P22064	homo sapien	
18	315.5	10.8	443	1	FBF14_HUMAN	O95967	homo sapien	
19	311	10.7	703	1	FBF1_HUMAN	P23142	homo sapien	
20	308.5	10.6	443	1	FBF14_CRIGR	O55058	cricetulus	
21	308	10.6	448	1	FBF15_HUMAN	O9ubx5	homo sapien	
22	308	10.6	705	1	FBF1_MOUSE	O08879	mus musculus	
23	307	10.5	684	1	FBF1_CHICK	O73775	gallus gall	
24	306.5	10.5	443	1	FBF4_MOUSE	Q9wvj9	mus musculus	
25	302	10.4	448	1	FBF5_MOUSE	Q9wvh9	mus musculus	
26	297	10.2	2531	1	NTC1_MOUSE	Q01705	mus musculus	
27	296	10.2	448	1	FBF5_RAT	Q9wvh8	rattus norv	
28	289.5	9.9	493	1	FBF13_HUMAN	Q12805	homo sapien	
29	284.5	9.8	2437	1	NOTC_BRARE	P46530	brachydanio	
30	282.5	9.7	2703	1	NOTC_DROME	P07207	drosophila	
31	280	9.6	2318	1	NTC3_MOUSE	Q61982	mus musculus	
32	278	9.5	2531	1	NTC1_RAT	O07008	rattus norv	
33	276.5	9.5	493	1	FBF13_RAT	O35568	rattus norv	

RT functionality by glycosyltransferase competition for serine474.";

RL Biochem. J. 295:131-140(1993).

CC [7]

CC STRUCTURE BY NMR OF 389-407.

CC MEDLINE-96007474; PubMed-7559494;

CC Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;

CC "The structure of a 19-residue fragment from the C-loop of the fourth

CC epidermal growth factor-like domain of thrombomodulin.";

CC J. Biol. Chem. 270:23366-23372(1995).

CC [9]

CC STRUCTURE BY NMR OF 364-407.

CC MEDLINE-96100636; PubMed-8528067;

CC Meininger D.P., Hunter M.J., Komives E.A.;

CC "Synthesis, activity, and preliminary structure of the fourth

CC EGF-like domain of thrombomodulin.";

CC Protein Sci. 4:1683-1695(1995).

CC [9]

CC STRUCTURE BY NMR OF 427-444.

CC MEDLINE-95034791; PubMed-7947766;

CC Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;

CC "Thrombin-bound structure of an EGF subdomain from human

CC thrombomodulin determined by transferred nuclear Overhauser

CC effects.";

CC Biochemistry 33:13553-13560(1994).

CC [10]

CC STRUCTURE BY NMR OF 427-444.

CC MEDLINE-96276211; PubMed-8745396;

CC Hrabal R., Komives E.A., Ni F.;

CC "Structural resiliency of an EGF-like subdomain bound to its target

CC protein, thrombin.";

CC Protein Sci. 5:195-203(1996).

CC [11]

CC STRUCTURE BY NMR OF 405-444.

CC MEDLINE-98035729; PubMed-9367781;

CC Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;

CC "Structure of the fifth EGF-like domain of thrombomodulin: an

CC EGF-like domain with a novel disulfide-bonding pattern.";

CC J. Mol. Biol. 273:913-926(1997).

CC [12]

CC VARIANT TED TYR-486.

CC PubMed-7811989;

CC Oehlin A.-K., Marlar R.A.;

CC "The first mutation identified in the thrombomodulin gene in a

CC 45-year-old man presenting with thromboembolic disease.";

CC Blood 85:330-336(1995).

CC [13]

CC VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.

CC PubMed-9198186;

CC Oehlin A.-K., Norlund L., Marlar R.A.;

CC "Thrombomodulin gene variations and thromboembolic disease.";

CC Thromb. Haemost. 78:396-400(1997).

CC [14]

CC VARIANT VAL-473.

CC PubMed-9157575;

CC Norlund L., Holm J., Zoller B., Oehlin A.-K.;

CC "A common thrombomodulin amino acid dimorphism is associated with

CC myocardial infarction.";

CC Thromb. Haemost. 77:248-251(1997).

CC [15]

CC VARIANT THR-43.

CC PubMed-9843165;

CC Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,

CC Stubbs P.J., Manger Cats V., Ireland H.;

CC "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,

CC and the risk of myocardial infarction in men.";

CC Thromb. Haemost. 80:743-748(1998).

CC [16]

CC VARIANT VAL-473.

CC PubMed-11245641;

CC Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,

CC Juneja H.;

CC "Thrombomodulin Ala455Val polymorphism and risk of coronary heart

CC disease.";

CC RT

Circulation 103:1386-1389(2001).

RL FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR

CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS

CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE

CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA

CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,

CC FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF

CC THROMBIN GENERATED.

CC SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING

CC THROMBOMODULIN.

CC POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED

CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).

CC DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,

CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE

CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE

CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE

CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.

CC SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.

CC DATABASE: NAME-PROW; NOTE-CD guide CD141 entry;

CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd141.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation--

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

EMBL; X05495; CAA29045.1; -

EMBL; M16552; AAB59508.1; -

EMBL; J02973; AAA61175.1; -

EMBL; D00210; BAA00149.1; -

EMBL; AL049651; CAB51954.1; -

PIR; A27073; A27073.

PIR; A28307; A28307.

PIR; A29680; A29680.

PDB; 1EGT; 15-NOV-95.

PDB; 1FGD; 20-JUN-96.

PDB; 1FGE; 20-JUN-96.

PDB; 1TMR; 08-JUN-95.

PDB; 1ZAO; 29-JAN-96.

PDB; 1ADX; 24-DEC-97.

PDB; 2ADX; 24-DEC-97.

GlycoSuiteDB; P07204; -

MIM; 188040; -

InterPro; IPR000152; Asx_hydroxyl.

InterPro; IPR000561; EGF-like.

InterPro; IPR001881; EGF_Ca.

InterPro; IPR001491; Thrombomoduln.

InterPro; IPR001304; lectin_c.

Pfam; PF00008; EGF; 5.

Pfam; PF00059; lectin_c; 1.

PRINTS; PR00907; THROMBOMODULN.

SMART; SM00034; CLECT; 1.

SMART; SM00179; EGF_CA; 1.

SMART; SM00001; EGF_like; 5.

PROSITE; PS00010; ASX_HYDROXYL; 2.

PROSITE; PS00022; EGF_1; FALSE_NEG.

PROSITE; PS01186; EGF_2; 2.

PROSITE; PS01187; EGF_CA; 2.

PROSITE; PS50041; C-TYPE LECTIN 2; 1.

Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;

KW Glycoprotein; Signal; EGF-like domain; Disease mutation; Polymorphism;

KW 3D-structure.

FT SIGNAL 1 21

FT CHAIN 22 575

FT DOMAIN 22 515

FT TRANSMEM 516 539

FT DOMAIN 540 575

FT DOMAIN 241 281

FT DOMAIN 284 324

FT POTENTIAL.

FT THROMBOMODULIN.

FT EXTRACELLULAR (POTENTIAL).

FT POTENTIAL.

FT CYTOPLASMIC (POTENTIAL).

FT EGF-LIKE 1.

FT EGF-LIKE 2.


```
FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT SIMILARITY)
SQ SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;

Query Match 65.3%; Score 1903; DB 1; Length 577;
Best Local Similarity 65.6%; Pred. No. 2.1e-118;
Matches 342; Conservative 46; Mismatches 125; Indels 8; Gaps 4;

QY 1 MLGVILGALALAGLPAPAEPOGSCQVEHDFCFALYPGAPFALNASQICDGLRHLM 60
DB 1 MLGIFLGLVAPASLGLSALAKLQPTGSCQVEHDFCFALFGPATFLDASACQLQGHLM 60
QY 61 TVRSSVAADVISLLNGDG-GVGRRRRLWIGLQPPGCGDKPKRLGRLGFQWVTGDNNTSY 119
DB 61 TVRSSVAADVISLLSQSSMDLGP--WIGLQLPGQCDPVLHGLPLRGFQWVTGDNNTSY 117
QY 120 SRWRLDLNGLAPLCGLVAVSAAETVPSEPIWEEOQCEVKADGFLCEPHFPATCRPLA 179
DB 118 SRWARPNDQAPLPGCLVTVSTATEAAGPEPAWEKPCETETQGFLCEFYFTASCRPLT 177
QY 180 VEP-GAAAAAVSYTYGPFPAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWARE 238
DB 178 VNTRPDPAAHISSTYNTYTPFGVGADFTLPVGSAAVEPLGLELIVCRAPPTSGEHWAVE 237
QY 239 APGAWDCSVENGCGEHAACNAIPGAPRCQCPAGAAQADGRSCSTASATQSCNDLCEHFVCP 298
DB 238 ATGAWNCSVENGCGEYLCNRSNEPRCLCPDRMDLQADGRSCARPVVSQCNELCEHFVCS 297
QY 299 NPQPGSYSCMCTGTGYRLADQHRCEVDVDCILEPSPCPCQRCVNTQGGFCHCYPNVDLV 358
DB 298 NAEVPGSYSCMCTGTGYRLADQHRCEVDVDCIKQGNPNQVNTKGGFCHCYVDGYELV 357
QY 359 DGEVCEVDPDFRANCYQCPNLTQSYLCVCAEGEAPIPHEPHRCOMFCNQACPADCD 418
DB 358 DGEVCEVDPDFRANCYQCPNLTQSYLCVCAEGEAPIPHEPHRCOMFCNQACPADCD 417
QY 419 PNTQASCEPEGYILDDGFTCTDIDECENGFGCGVCHNLPDGFCEICGSDSALARHIGT 478
DB 418 PNSTVCECEPEGYILDDGFTCTDIDECENGFGCGVCHNLPDGFCEICGSDSALARHIGT 477
QY 479 DCDGSKV---DGDGDSGSEPPPTFTGTLTPPAVLVHSG 516
DB 478 DCDPIPVREDTKEEGSGEPPVPTPGSPTGPPSARPVHSG 518

RESULT 3
TRBM_BOVIN STANDARD; PRT; 356 AA.
AC P06579;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE THROMBOMODULIN (FETOMODULIN) (TM) (FRAGMENT).
GN THBD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067408; PubMed=3024152;
RA Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;
RT "Characterization of a thrombomodulin cDNA reveals structural
RT similarity to the low density lipoprotein receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).
CC -!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
CC SCISSIONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
```


or send an email to license@isb-sib.ch).

CC EMBL; L28748; AAA74122.1; --
DR HSP; P35555; IAPJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002212; TB.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBLD.
DR SMART; SM00179; EGF_Ca; 42.
DR SMART; SM00001; EGF_Ca; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_Ca; 45.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 330 390 TGFBB 1.
FT DOMAIN 392 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 654 722 TGFBB 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 952 1027 TGFBB 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1528 1605 TGFBB 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1689 1765 TGFBB 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2055 2126 TGFBB 6.
FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2333 2401 TGFBB 7.
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.

2524 DOMAIN
2567 FT
2607 FT
2647 FT
2648 FT
85 FT
94 FT
2566 EGF-LIKE 44, CALCIUM-BINDING.
2606 EGF-LIKE 45, CALCIUM-BINDING.
2647 EGF-LIKE 46, CALCIUM-BINDING.
2687 EGF-LIKE 47, CALCIUM-BINDING.
94 BY SIMILARITY.
100 BY SIMILARITY.
111 BY SIMILARITY.
129 BY SIMILARITY.
134 BY SIMILARITY.
145 BY SIMILARITY.
160 BY SIMILARITY.
166 BY SIMILARITY.
177 BY SIMILARITY.
262 BY SIMILARITY.
271 BY SIMILARITY.
286 BY SIMILARITY.
292 BY SIMILARITY.
304 BY SIMILARITY.
313 BY SIMILARITY.
328 BY SIMILARITY.
455 BY SIMILARITY.
474 BY SIMILARITY.
488 BY SIMILARITY.
504 BY SIMILARITY.
513 BY SIMILARITY.
528 BY SIMILARITY.
546 BY SIMILARITY.
555 BY SIMILARITY.
570 BY SIMILARITY.
587 BY SIMILARITY.
596 BY SIMILARITY.
611 BY SIMILARITY.
628 BY SIMILARITY.
637 BY SIMILARITY.
652 BY SIMILARITY.
739 BY SIMILARITY.
748 BY SIMILARITY.
763 BY SIMILARITY.
781 BY SIMILARITY.
790 BY SIMILARITY.
805 BY SIMILARITY.
821 BY SIMILARITY.
830 BY SIMILARITY.
845 BY SIMILARITY.
926 BY SIMILARITY.
935 BY SIMILARITY.
937 BY SIMILARITY.
950 BY SIMILARITY.
1044 BY SIMILARITY.
1053 BY SIMILARITY.
1068 BY SIMILARITY.
1086 BY SIMILARITY.
1095 BY SIMILARITY.
1111 BY SIMILARITY.
1129 BY SIMILARITY.
1138 BY SIMILARITY.
1153 BY SIMILARITY.
1171 BY SIMILARITY.
1180 BY SIMILARITY.
1195 BY SIMILARITY.
1212 BY SIMILARITY.
1221 BY SIMILARITY.
1236 BY SIMILARITY.
1254 BY SIMILARITY.
1263 BY SIMILARITY.
1278 BY SIMILARITY.
1296 BY SIMILARITY.
1305 BY SIMILARITY.
1320 BY SIMILARITY.
1339 BY SIMILARITY.
1348 BY SIMILARITY.
1361 BY SIMILARITY.
1380 BY SIMILARITY.
1389 BY SIMILARITY.
1402 BY SIMILARITY.
1391 FT
1361 FT
1380 FT
1391 FT
1402 FT


```
FT DISULFID 928 941 BY SIMILARITY.
FT DISULFID 947 956 BY SIMILARITY.
FT DISULFID 952 965 BY SIMILARITY.
FT DISULFID 967 980 BY SIMILARITY.
FT DISULFID 986 998 BY SIMILARITY.
FT DISULFID 994 1007 BY SIMILARITY.
FT DISULFID 1009 1023 BY SIMILARITY.
FT DISULFID 1029 1042 BY SIMILARITY.
FT DISULFID 1036 1051 BY SIMILARITY.
FT DISULFID 1056 1068 BY SIMILARITY.
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F9EC5D CRC64;

Query Match 12.6%; Score 367.5; DB 1; Length 1184;
Best Local Similarity 30.0%; Pred. No. 3.8e-17;
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

QY 154 EQOCEVKAD-GFCEPHFPATC-----RPLAV-----EPGAAAASVITYGTPFAARG 201
Db 531 EQSCESNPNGYPCN-HVMLSCEGEELTVPEVRRPPEPAAPRRVS-----EAEM 582
QY 202 ADFALPVGSSAAV---APLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258
Db 583 AGRALSIGTEAEALPNSLPGDDQDECLLPGEL-----COHLICIN 622
QY 259 IPGAPRCQCPAGAAQLQADGRSC-----TSAT-----QS 287
Db 623 TVGSYHCACFPFSLQDDGRTCRPEGHPPQPEAPQEPALKSEFQVASFSTIPLPLPQNT 682
QY 288 CND--LCEHFCVNPDPQGSVSCMCETGYRLAADOHRCEVDVDDCILEPSPCP--ORCVNT 343
Db 683 CKDNGPKQVC---STVGSAICSCFPFYALMADGVSCDINECVTLHTCSRGEHCVNT 739
QY 344 QGFECH---CYPNYDLVDCVEPVDPCFRANCEYQCQPLNQTSLYCV-----389
Db 740 LGSFHCYKALTCEPGYALKDGEK-EDVDEC--AMGTHTCQP-----GFLCQNTKGSFYCQA 792
QY 390 ---CAEGFAPLPH-----BPHRCQFCNQTACPADCPDNPNTQASCEGPEGV-I 432
Db 793 RQRCMDGLQDPENGVNDINETSLEPCRFSGFSCINTVGSYTCQNPPLI---CARGIHA 849
QY 433 LDDGFICTIDECENGFCGSG---VCHNLPGTFECIC 466
Db 850 SDDGAKVDVNECETGVHRCGEGQVCHNLPGSYRDC 886

RESULT 7
FBNI_PIG
ID FBNI_PIG STANDARD; PRT; 2871 AA.
AC Q9TV36;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FBIRILLIN 1 PRECURSOR.
GN FBNI.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Lung;
RX MEDLINE=99156858; PubMed=10036187;
RA Bieri N.J., Eldadah Z.A., Moore C.S., Stetten G., Spencer F.,
RA Dietz H.C.;
RT "Revised genomic organization of FBNI and significance for regulated
RT gene expression.";
RL Genomics 56:70-77(1999).
CC -I- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC -I- THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
```

```
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -I- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS.
CC -I- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF073800; AAD50328.1; -.
CC HSSP; P35555; IAPJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002212; TB.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00008; EGF; 45.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBLLOOD.
DR SMART; SM00179; EGF_CA; 40.
DR SMART; SM00001; EGF_Like; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 41.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 41.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 330 390 TGFBP 1.
FT DOMAIN 392 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 654 722 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 952 1027 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1528 1605 TGFBP 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1689 1765 TGFBP 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
```

```
FT DOMAIN 1849 EGF-LIKE 31, CALCIUM-BINDING. 1890 BY SIMILARITY. 1236
FT DOMAIN 1891 EGF-LIKE 32, CALCIUM-BINDING. 1929 BY SIMILARITY. 1254
FT DOMAIN 1930 EGF-LIKE 33, CALCIUM-BINDING. 1972 BY SIMILARITY. 1263
FT DOMAIN 1973 EGF-LIKE 34, CALCIUM-BINDING. 2012 BY SIMILARITY. 1278
FT DOMAIN 2013 EGF-LIKE 35, CALCIUM-BINDING. 2054 BY SIMILARITY. 1296
FT REPEAT 2055 TGFBP 6. 2126 BY SIMILARITY. 1305
FT DOMAIN 2127 EGF-LIKE 36, CALCIUM-BINDING. 2165 BY SIMILARITY. 1320
FT DOMAIN 2166 EGF-LIKE 37, CALCIUM-BINDING. 2205 BY SIMILARITY. 1339
FT DOMAIN 2206 EGF-LIKE 38, CALCIUM-BINDING. 2246 BY SIMILARITY. 1348
FT DOMAIN 2247 EGF-LIKE 39, CALCIUM-BINDING. 2291 BY SIMILARITY. 1361
FT DOMAIN 2291 EGF-LIKE 40, CALCIUM-BINDING. 2332 BY SIMILARITY. 1380
FT REPEAT 2333 TGFBP 7. 2401 BY SIMILARITY. 1391
FT DOMAIN 2402 EGF-LIKE 41, CALCIUM-BINDING. 2443 BY SIMILARITY. 1402
FT DOMAIN 2444 EGF-LIKE 42, CALCIUM-BINDING. 2484 BY SIMILARITY. 1420
FT DOMAIN 2485 EGF-LIKE 43, CALCIUM-BINDING. 2523 BY SIMILARITY. 1429
FT DOMAIN 2524 EGF-LIKE 44, CALCIUM-BINDING. 2566 BY SIMILARITY. 1461
FT DOMAIN 2567 EGF-LIKE 45, CALCIUM-BINDING. 2606 BY SIMILARITY. 1470
FT DOMAIN 2607 EGF-LIKE 46, CALCIUM-BINDING. 2647 BY SIMILARITY. 1485
FT DOMAIN 2648 EGF-LIKE 47, CALCIUM-BINDING. 2687 BY SIMILARITY. 1511
FT DISULFID 85 94 BY SIMILARITY. 1526
FT DISULFID 89 100 BY SIMILARITY. 1622
FT DISULFID 102 111 BY SIMILARITY. 1631
FT DISULFID 119 129 BY SIMILARITY. 1646
FT DISULFID 123 134 BY SIMILARITY. 1663
FT DISULFID 136 145 BY SIMILARITY. 1672
FT DISULFID 150 160 BY SIMILARITY. 1674
FT DISULFID 154 166 BY SIMILARITY. 1687
FT DISULFID 168 177 BY SIMILARITY. 1770
FT DISULFID 177 188 BY SIMILARITY. 1782
FT DISULFID 250 262 BY SIMILARITY. 1791
FT DISULFID 257 271 BY SIMILARITY. 1806
FT DISULFID 273 286 BY SIMILARITY. 1824
FT DISULFID 292 304 BY SIMILARITY. 1833
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.

Query Match 12.6%; Score 366; DB 1; Length 2871;
Best Local Similarity 30.6%; Pred. No. 1.le-16;
Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGAAQADORSCTASATQSCNDLCEHFCVNP--- 300
DB 1200 ECSIMNGGCTFCNTSEGSYCQPGFALPDQRST-----DIDE--CEDNPNIC 1249

QY 301 -----DQPGSYSCMCETGYRLAADQHRCEDDVDCILEPSPC-PQRCVNTGGFECHCYP 353
DB 1250 DGGCTNIPGEYRCLCYDGFWASEDMKTCVDVNECDLNPICLSGTCENTKGSFICHDM 1309

QY 354 NYDLVDGE--CVEPVDP--FRANCEYQCOPLNOT-SYLCVCARGFA-----PIP 398
DB 1310 GYSGKKGKGTCTD-INECEIGAHNCDRHAHVCTNTAGSFNCSPGWIGDGKICTDLDECS 1368

QY 399 HEPHRCQFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDEC-BN----- 447
DB 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCADLDECSENVKLCGNVQ 1419

QY 448 -----GGF-----CS-----GVCHNLPGTFFECICGPDALA 473
DB 1420 CLYAPGGYHCEYDMGFVPSADRSKSCVDSDECSLPNICVFTCHNLPLGLFRCEC----- 1472

QY 474 RHIGTDCDCKVDGDSGSGEPSPS-----PTPGS-----TLTPPAVGLV 513
DB 1473 -EIGYELDRSGGCTDVNECLEPPTCISGNCVNTPGSYTCVCPDPFELNTRVGCV 1527

RESULT 8
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBRILLIN 2 PRECURSOR.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

DR	PRINTS; PR00010; EGF-BLOOD.
DR	DR SMART; SM00179; EGF_CA; 43.
DR	DR SMART; SM00001; EGF_Like; 3.
DR	DR PROSITE; PS00010; ASX-HYDROXYL; 43.
DR	DR PROSITE; PS00022; EGF_L1; 2.
DR	DR PROSITE; PS01186; EGF_L2; 37.
DR	DR PROSITE; PS01187; EGF_CA; 43.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain.
KW	Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
FT	SIGNAL 1 28
FT	CHAIN 29 2911
FT	DOMAIN 111 142
FT	DOMAIN 145 176
FT	DOMAIN 176 207
FT	DOMAIN 275 316
FT	DOMAIN 317 358
FT	REPEAT 359 425
FT	DOMAIN 493 533
FT	DOMAIN 534 573
FT	DOMAIN 574 615
FT	DOMAIN 616 656
FT	DOMAIN 657 697
FT	REPEAT 698 766
FT	DOMAIN 767 808
FT	DOMAIN 809 850
FT	DOMAIN 851 890
FT	DOMAIN 954 995
FT	REPEAT 996 1071
FT	DOMAIN 1072 1113
FT	DOMAIN 1114 1156
FT	DOMAIN 1157 1198
FT	DOMAIN 1199 1240
FT	DOMAIN 1241 1281
FT	DOMAIN 1282 1323
FT	DOMAIN 1324 1365
FT	DOMAIN 1366 1406
FT	DOMAIN 1407 1447
FT	DOMAIN 1448 1489
FT	DOMAIN 1490 1530
ET	DOMAIN 1531 1571
FT	REPEAT 1572 1648
FT	DOMAIN 1649 1690
FT	DOMAIN 1691 1732
FT	REPEAT 1733 1806
FT	DOMAIN 1807 1848
ET	DOMAIN 1849 1890
FT	DOMAIN 1891 1932
FT	DOMAIN 1933 1971
FT	DOMAIN 1972 2014
FT	DOMAIN 2015 2054
FT	DOMAIN 2055 2096
FT	REPEAT 2097 2169
FT	DOMAIN 2170 2211
FT	DOMAIN 2212 2251
FT	DOMAIN 2252 2292
FT	DOMAIN 2293 2336
FT	DOMAIN 2337 2378
FT	REPEAT 2379 2447
FT	DOMAIN 2448 2489
FT	DOMAIN 2490 2530
FT	DOMAIN 2531 2569
FT	DOMAIN 2570 2612
FT	DOMAIN 2613 2652
FT	DOMAIN 2653 2693
FT	DOMAIN 2694 2733
FT	DISULFID 115 124
FT	DISULFID 119 130
FT	DISULFID 132 141
ET	DISULFID 149 159
FT	DISULFID 153 164
FT	DISULFID 166 175
FT	DISULFID 180 190
FT	DISULFID 184 195

FT	DISULFID	197	206	BY SIMILARITY.
FT	DISULFID	279	291	BY SIMILARITY.
FT	DISULFID	286	300	BY SIMILARITY.
FT	DISULFID	302	315	BY SIMILARITY.
FT	DISULFID	321	333	BY SIMILARITY.
FT	DISULFID	328	342	BY SIMILARITY.
FT	DISULFID	344	357	BY SIMILARITY.
FT	DISULFID	497	509	BY SIMILARITY.
FT	DISULFID	504	518	BY SIMILARITY.
FT	DISULFID	520	532	BY SIMILARITY.
FT	DISULFID	538	548	BY SIMILARITY.
FT	DISULFID	543	557	BY SIMILARITY.
FT	DISULFID	559	572	BY SIMILARITY.
FT	DISULFID	578	590	BY SIMILARITY.
FT	DISULFID	585	599	BY SIMILARITY.
FT	DISULFID	601	614	BY SIMILARITY.
FT	DISULFID	620	631	BY SIMILARITY.
FT	DISULFID	626	640	BY SIMILARITY.
FT	DISULFID	642	655	BY SIMILARITY.
FT	DISULFID	661	672	BY SIMILARITY.
FT	DISULFID	667	681	BY SIMILARITY.
FT	DISULFID	683	696	BY SIMILARITY.
FT	DISULFID	771	783	BY SIMILARITY.
FT	DISULFID	778	792	BY SIMILARITY.
FT	DISULFID	794	807	BY SIMILARITY.
FT	DISULFID	813	825	BY SIMILARITY.
FT	DISULFID	820	834	BY SIMILARITY.
FT	DISULFID	836	849	BY SIMILARITY.
FT	DISULFID	855	865	BY SIMILARITY.
FT	DISULFID	860	874	BY SIMILARITY.
FT	DISULFID	876	889	BY SIMILARITY.
FT	DISULFID	958	970	BY SIMILARITY.
FT	DISULFID	965	979	BY SIMILARITY.
FT	DISULFID	981	994	BY SIMILARITY.
FT	DISULFID	1076	1088	BY SIMILARITY.
FT	DISULFID	1083	1097	BY SIMILARITY.
FT	DISULFID	1099	1112	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1139	BY SIMILARITY.
FT	DISULFID	1141	1155	BY SIMILARITY.
FT	DISULFID	1161	1173	BY SIMILARITY.
FT	DISULFID	1168	1182	BY SIMILARITY.
FT	DISULFID	1184	1197	BY SIMILARITY.
FT	DISULFID	1203	1215	BY SIMILARITY.
FT	DISULFID	1210	1224	BY SIMILARITY.
Query Match 12.6%; Score 366; DB 1; Length 2911;				
Best Local Similarity 33.1%; Pred. No. 1.le-16;				
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps				
QY	244	DCSVENGGEHACNATPGAPRCQCPAGALOADGRCSFTASATQSCNDLCEHFCVNPDP--	301	
DB	1244	ECIMNGGCDTQCTNSEGSEYSCSEGEALMPDGRSCA-----DIDE--CENNPDIC	1293	
QY	302	-----QPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCP--QRCVNTQGGFECHCYP	353	
DB	1294	DGGQCTNIPGEYRCLCYDGFMA5MDMKTCIDVNECDLNSNICMFEGCENTKGSFICHQCL	1353	
QY	354	NYDLVGE--CUEVPDPC--FRANCEYQOPLN-OTSYLVCVACGFAPIPH-----EPHR	403	
DB	1354	GYSVKRGTTGCTD-VDECEIGAHDMDHASCLINIPGSPKCSREGW--IGNIKCIDLDE	1410	
QY	404	QMFECNOTACPADCDNTQAS--CPCPEGYILDDGFICTDIDE-----CENG-----	448	
DB	1411	CSNGTHQCSINAQC-VNTPG5YRCACSEGF-TGDGFTCDSDVDECAENINLCENGQCLNVP	1468	
QY	449	-----GF-----CSGVCNHLPGTFECICGPD5ALARHIG--	477	
DB	1469	GAYRCECEMGFTPASDRSCQDIDBCSFQNICVSGTCNNLFGMFHCICDDGYELDRITGN	1528	
QY	478	-TDCD	481	
DB	1529	CTDID	1533	


```

Db 1310 GYSGKKGKGTCTD-INECEIGHNCKGRHAVCTNTAGSFKSCSPGWIGDGIKCTDLDECS 1368
Qy 399 HEPHRCMFCNQTCAPADCDPNTQAS--CECEPEGYILDGFCIDIDBC-EN----- 447
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1369 NGTHCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDCSENLCGNQ 1419
Qy 448 -----GGF-----CS-----GVCNLPGTFCICGPDLSALA 473
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1420 CLNAPGGYRCEDMGFVPSADGKACEDIDECSLPNICVFGTCHNLPGLFCECEIGVELD 1479
Qy 474 RHIG-----TDCDSKVDGSDSGSGEPPTPEGS-----TLTPPAVGLV 513
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1480 RSGNCTDVNECLDPTTCISGCVN-----TPGSYICDCPPDFELNTRVGC 1527

RESULT 10
FBNL_MOUSE
ID FBNL_MOUSE STANDARD; PRT; 2871 AA.
AC Q61554; Q60826;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBRILLIN 1 PRECURSOR.
GN FBN1 OR FBN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95130561; PubMed=7829516;
RA Yin W., Germiller J., Sanguinetti C., Smiley E., Pangillan T.,
RA Pereira L., Ramirez F., Bonadio J.;
RT "Primary structure and developmental expression of Fbn-1, the mouse
RL J. Biol. Chem. 270:1798-1806(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney;
RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RA Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- PPM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29454; AAA56840.1; -.
CC EMBL; U22493; AAA64217.1; -.
CC HSSP; P35555; IAPJ.
CC MGD; MGI:95489; Fbnl.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR002212; TB.
CC Pfam; PF00008; EGF; 46.
CC Pfam; PF00683; TB; 9.
CC SMART; SM00179; EGF_CA; 42.
CC SMART; SM00001; EGF_like; 4.
CC PROSITE; PS00010; ASX-HYDROXYL; 43.
CC PROSITE; PS00022; EGF_L1; 2.

```

```

DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 45.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT DOMAIN 330 401 TGFBP 1.
FT DOMAIN 402 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT DOMAIN 656 721 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT DOMAIN 952 1018 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT DOMAIN 1528 1599 TGFBP 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT DOMAIN 1689 1758 TGFBP 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT DOMAIN 2055 2121 TGFBP 6.
FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT DOMAIN 2333 2400 TGFBP 7.
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.

```

FT	DISULFID	292	304	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.
FT	DISULFID	315	328	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.
FT	DISULFID	499	513	BY SIMILARITY.
FT	DISULFID	515	528	BY SIMILARITY.
FT	DISULFID	534	546	BY SIMILARITY.
FT	DISULFID	541	555	BY SIMILARITY.
FT	DISULFID	557	570	BY SIMILARITY.
FT	DISULFID	576	587	BY SIMILARITY.
FT	DISULFID	582	596	BY SIMILARITY.
FT	DISULFID	598	611	BY SIMILARITY.
FT	DISULFID	617	628	BY SIMILARITY.
FT	DISULFID	623	637	BY SIMILARITY.
FT	DISULFID	639	652	BY SIMILARITY.
FT	DISULFID	727	739	BY SIMILARITY.
FT	DISULFID	734	748	BY SIMILARITY.
FT	DISULFID	750	763	BY SIMILARITY.
FT	DISULFID	769	781	BY SIMILARITY.
FT	DISULFID	776	790	BY SIMILARITY.
FT	DISULFID	792	805	BY SIMILARITY.
FT	DISULFID	811	821	BY SIMILARITY.
FT	DISULFID	816	830	BY SIMILARITY.
FT	DISULFID	832	845	BY SIMILARITY.
FT	DISULFID	914	926	BY SIMILARITY.
FT	DISULFID	921	935	BY SIMILARITY.
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1111	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.
FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	1658	1672	BY SIMILARITY.
Query Match				12.4%; Score 362; DB 1; Length 2871;

Best Local Similarity 29.4%; Pred. No. 1.9e-16;			
Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;			
QY	212	SAAVAPLGL-----QLMCTAPPVQGVHAREAPGAWDSVNGGCEHACNAIPGAPRC	265
Db	1162	SANLCPHGRVNLICKYQCACPNGYHPTDRLFCVDIDECIMNGGCEFTCTNSDGSYEC	1221
QY	266	QCPAGAAALQADGSCSTASATQSCNDLCEHFCVPNP-----DOPGSYSCMCTGYRL	316
Db	1222	SCQPGFALMPDQRST-----DIDQ--CEDPNICDGGQCTNIPGEYRCLCYDGFMA	1271
QY	317	AADQHRCEVDVDCILEPSPC-PQRCVNTGGFCYCHYPNYDLVDGE--CVEPVDPC--FR	371
Db	1272	SEDMTKVDVNECDLNPICLSGTCENTKGSFICHDMGYSKKKGTGCTD--INECEIGA	1330
QY	372	ANCEYQCQPLNOT--SYLCVCAEGFA-----PIPEPHRCQMFQNOTACPADCPN	420
Db	1331	HNCGRHACTNTAGSFKSCSPGWIGDKIKTDLDECSTNGTHWCSOH-----ADC-KN	1382
QY	421	TOAS--CECEGYILDDGFICTDIDEC-EN-----GGF-----	450
Db	1383	TMGSYRCLCKDGY-TGDGFTCTDLDECSENLNLCGNGQCLNAPGGYRCDCMGFVPSADG	1441
QY	451	-----CS-----GVCHNLPGTFECICGPDSSALARHIG-----TDCDSK	484
Db	1442	KACEDIDECLEPNICVFGTCHNLPLFRCEIGYELDRSGNCTDVNECLDPTTCISGN	1501
QY	485	VDGDSGSGSEPPSPTPGS-----TLTPPAVGLV	513
Db	1502	CVN-----TPGSYTCDCSPDFELNTRVGCV	1527
RESULT 11			
FBL2_MOUSE STANDARD; PRT: 1221 AA.			
AC	P37889; Q9WUI2;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-OCT-1994 (Rel. 30, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	FIBULIN-2 PRECURSOR.		
GN	FBLN2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.		
RC	TISSUE=Fibroblast;		
RX	MEDLINE=94064787; PubMed=8245130;		
RA	Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;		
RT	"Structure and expression of fibulin-2, a novel extracellular matrix		
RT	protein with multiple EGF-like repeats and consensus motifs for		
RT	calcium binding.";		
RL	J. Cell Biol. 123:1269-1277(1993).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99337686; PubMed=10406956;		
RA	Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;		
RT	"Mouse fibulin-2 gene. Complete exon-intron organization and promoter		
RT	characterization.";		
RL	Eur. J. Biochem. 263:471-477(1999).		
CC	-1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS		
CC	-1- CALCIUM DEPENDENT.		
CC	-1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.		
CC	-1- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS OF FIBULIN-2 DUE TO THE		
CC	PRESENCE OR ABSENCE OF A SINGLE EGF-LIKE (3) REPEAT MAY ARISE BY		
CC	ALTERNATIVE SPLICING.		
CC	-1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER		
CC	CONNECTIVE TISSUES.		
CC	-1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.		

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X75285; CAA53040.1; .
 DR EMBL; AF135253; AAD34456.1; JOINED.
 DR EMBL; AF135239; AAD34456.1; JOINED.
 DR EMBL; AF135240; AAD34456.1; JOINED.
 DR EMBL; AF135241; AAD34456.1; JOINED.
 DR EMBL; AF135242; AAD34456.1; JOINED.
 DR EMBL; AF135243; AAD34456.1; JOINED.
 DR EMBL; AF135244; AAD34456.1; JOINED.
 DR EMBL; AF135245; AAD34456.1; JOINED.
 DR EMBL; AF135246; AAD34456.1; JOINED.
 DR EMBL; AF135247; AAD34456.1; JOINED.
 DR EMBL; AF135248; AAD34456.1; JOINED.
 DR EMBL; AF135249; AAD34456.1; JOINED.
 DR EMBL; AF135250; AAD34456.1; JOINED.
 DR EMBL; AF135251; AAD34456.1; JOINED.
 DR EMBL; AF135252; AAD34456.1; JOINED.
 DR PIR; A49457; A49457.
 DR HSP; P07204; IrGD.
 DR MGD; MGI:95488; Fbln2.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR Pfam; PF01821; ANATO; 2.
 DR Pfam; PF00008; EGF; 6.
 DR SMART; SM00104; ANATO; 3.
 DR SMART; SM00179; EGF_Ca; 9.
 DR SMART; SM00001; EGF_like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 5.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_Ca; 10.
 KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
 KW Calcium-binding; Alternative splicing; Repeat.
 FT SIGNAL 1 26
 FT CHAIN 27 1221
 FT DOMAIN 27 434
 FT DOMAIN 27 176
 FT DOMAIN 177 434
 FT DOMAIN 435 477
 FT DOMAIN 478 510
 FT DOMAIN 511 543
 FT DOMAIN 594 635
 FT DOMAIN 669 708
 FT DOMAIN 709 755
 FT DOMAIN 756 800
 FT DOMAIN 801 846
 FT DOMAIN 847 894
 FT DOMAIN 895 937
 FT DOMAIN 938 979
 FT DOMAIN 980 1018
 FT DOMAIN 1019 1061
 FT DOMAIN 1062 1106
 FT DOMAIN 1111 1221
 FT SITE 421 423
 FT DISULFID 435 462
 FT DISULFID 436 469
 FT DISULFID 449 470
 FT DISULFID 479 508
 FT DISULFID 492 509
 FT DISULFID 511 535
 FT DISULFID 512 542
 FT DISULFID 525 543
 N.
 SUBDOMAIN NA (CYS-RICH).
 SUBDOMAIN NB (CYS-FREE).
 ANAPHYLATOXIN-LIKE 1.
 ANAPHYLATOXIN-LIKE 2.
 ANAPHYLATOXIN-LIKE 3.
 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 2.
 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 DOMAIN III.
 CELL ATTACHMENT SITE (POTENTIAL).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

FT DISULFID 598 610 BY SIMILARITY.
 FT DISULFID 606 619 BY SIMILARITY.
 FT DISULFID 621 634 BY SIMILARITY.
 FT DISULFID 673 683 BY SIMILARITY.
 FT DISULFID 679 692 BY SIMILARITY.
 FT DISULFID 694 707 BY SIMILARITY.
 FT DISULFID 713 726 BY SIMILARITY.
 FT DISULFID 720 735 BY SIMILARITY.
 FT DISULFID 742 754 BY SIMILARITY.
 FT DISULFID 805 818 BY SIMILARITY.
 FT DISULFID 812 827 BY SIMILARITY.
 FT DISULFID 833 845 BY SIMILARITY.
 FT DISULFID 899 912 BY SIMILARITY.
 FT DISULFID 923 936 BY SIMILARITY.
 FT DISULFID 942 954 BY SIMILARITY.
 FT DISULFID 950 963 BY SIMILARITY.
 FT DISULFID 965 978 BY SIMILARITY.
 FT DISULFID 984 993 BY SIMILARITY.
 FT DISULFID 1004 1017 BY SIMILARITY.
 FT DISULFID 1023 1035 BY SIMILARITY.
 FT DISULFID 1031 1044 BY SIMILARITY.
 FT DISULFID 1046 1060 BY SIMILARITY.
 FT DISULFID 1066 1079 BY SIMILARITY.
 FT DISULFID 1073 1088 BY SIMILARITY.
 FT DISULFID 1093 1105 BY SIMILARITY.
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 709 755 MISSING (IN ISOFORM EGF3-LESS).
 FT CONFLICT 140 159 HSGRKYAAGHTVHLSSCRAC -> TVAVSICWYRPPLLP
 FT CONFLICT 348 348 S -> L (IN REF. 2).
 FT CONFLICT 507 507 Q -> QQ (IN REF. 2).
 FT CONFLICT 1102 1102 Q -> E (IN REF. 2).
 SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;
 Query Match 12.2%; Score 357; DB 1; Length 1221;
 Best Local Similarity 29.5%; Pred. No. 1.9e-16;
 Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;
 QY 154 EQQCEVKAD-GFLCEPHFPATC-----RPLAV-----EPGAAAAVS-----IT 192
 DB 521 EGQSCSNPNLGYPCN-HVMLSCEGEPLIVPEVRRPPEPEAAPRRVSEMEMASREALS 579
 QY 193 YGT----PFAARGADFO---ALP-----VGSSAAVAPLGLQLM-----CTAPPG 229
 DB 580 LGTEALPNSLPGDDQDECLMLPGLCQLHCINTVGSYRCACFFGPELQGGRTCRDRG 639
 QY 230 AVQGHWARE-APGAWDCSV-----ENGGEHACNAIPGAPRCQCPAGAA 272
 DB 640 APQLDTARESPRESAQVSEPTIPLVPQNTCKDNGPCQVCRVYVGDTCAMCSFCGYA 699
 QY 273 LQADGRSC-----TASATQSCNDLCEHFCVNPDPQPSYSC-----MCETYRLAAQDHR 322
 DB 700 IMADGVSCEDQDECLMGTIDCS--WKQFCV---NTLSFYCVNHTVLCAEGYILNA-HRK 753
 QY 323 CEDVDDCILEPSPC--PQRCVNTGGFECH-----CYPNYDLVDGECVPEVDPDPCFRA--NC 374
 DB 754 CVDINECVTLHTCTRAEHCNVNTPGSCYKALTCFPGYVLTGDECTD-VDECVTGTGHC 812
 QY 375 E--YQCOPLNOTSYLCV---CAEGFAPIPH-----EPHRCQMFQNOTACPA 415
 DB 813 QAGFSCQN-TKGSFYCOARQRCMDGFLQDPGNCVDINECTSLLEPCRSFSCINTVGSY 871
 QY 416 DCDPNTQASCECPGEGY-ILDDGFICTDIDECENGFCGSG---VCHNLPGTPECICGP 468
 DB 872 TCQRNPLV---CGRGYHANEESGSCVDNCECTGVHRCGECQLCYNLPGSYRCDCPK 925

FT	DOMAIN	320	360	EGF-LIKE 3.
FT	DOMAIN	361	401	EGF-LIKE 4.
FT	DOMAIN	402	442	EGF-LIKE 5.
FT	DOMAIN	443	483	EGF-LIKE 6.
FT	DOMAIN	484	524	EGF-LIKE 7.
FT	DOMAIN	525	565	EGF-LIKE 8.
FT	DOMAIN	566	606	EGF-LIKE 9.
FT	DOMAIN	607	647	EGF-LIKE 10.
FT	DOMAIN	655	830	WFA 2.
FT	DOMAIN	917	955	COILED COIL (POTENTIAL).
FT	DISULFID	242	253	BY SIMILARITY.
FT	DISULFID	249	262	BY SIMILARITY.
FT	DISULFID	264	277	BY SIMILARITY.
FT	DISULFID	283	294	BY SIMILARITY.
FT	DISULFID	290	303	BY SIMILARITY.
FT	DISULFID	305	318	BY SIMILARITY.
FT	DISULFID	324	335	BY SIMILARITY.
FT	DISULFID	331	344	BY SIMILARITY.
FT	DISULFID	346	359	BY SIMILARITY.
FT	DISULFID	365	376	BY SIMILARITY.
FT	DISULFID	372	385	BY SIMILARITY.
FT	DISULFID	387	400	BY SIMILARITY.
FT	DISULFID	406	417	BY SIMILARITY.
FT	DISULFID	413	426	BY SIMILARITY.
FT	DISULFID	428	441	BY SIMILARITY.
FT	DISULFID	447	458	BY SIMILARITY.
FT	DISULFID	454	467	BY SIMILARITY.
FT	DISULFID	469	482	BY SIMILARITY.
FT	DISULFID	488	499	BY SIMILARITY.
FT	DISULFID	495	508	BY SIMILARITY.
FT	DISULFID	510	523	BY SIMILARITY.
FT	DISULFID	529	540	BY SIMILARITY.
FT	DISULFID	536	549	BY SIMILARITY.
FT	DISULFID	551	564	BY SIMILARITY.
FT	DISULFID	570	581	BY SIMILARITY.
FT	DISULFID	577	590	BY SIMILARITY.
FT	DISULFID	592	605	BY SIMILARITY.
FT	DISULFID	611	622	BY SIMILARITY.
FT	DISULFID	618	631	BY SIMILARITY.
FT	DISULFID	633	646	BY SIMILARITY.
FT	CARBOHYD	221	221	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC	861	879	MISSING (IN SHORT ISOFORM).
FT	CONFLICT	356	356	K -> E (IN REF. 3).
FT	CONFLICT	594	594	V -> E (IN REF. 3).
FT	CONFLICT	644	644	R -> G (IN REF. 2).
FT	CONFLICT	755	755	F -> L (IN REF. 3).
FT	CONFLICT	935	935	L -> F (IN REF. 2).
FT	SEQUENCE	956 AA;	106840 MW;	826B7F347178FC80 CRC64;
QY	245	CSYVNGGCEHACNAIPCAPRCQCPAGAAQADGSRCTA-SATQSCNDLCEHFCVNPNDQP	303	
Db	283	CAMEDHNCCEGLCVNVPGSFVCQCYSGVALADGKRCVADVASCENHGCEHCY-NAD--	339	
QY	304	GSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDL-VDGEC	362	
Db	340	GSYLQCQHEGALNPDKTKTIDYCASSNHGCOHCVCVNTDDSYSCHLKGFTLNPDKKT	399	
QY	363	VEPVDPCF--RANCEYQCQLNQTSYLCVCAEGFAPIPH-----EPHRCQMFC	408	
Db	400	RRINYCALNKPGEHCVCNMEE--SYRCRCHRGYTLDPNGKTCRSRVDHCAQDDHGCEQLC	458	
QY	409	NQTACPADCPDNTQAS--CECPGYILDDGF-ICTDIDEC---ENGFCSCVCHNLPGTF	462	
Db	459	-----INTDESFVCQCSGEGFLNEDLKTCSRVDYCLLSDHG--CEYSCVNMDRSF	506	
QY	463	ECICGPDPSALARHIGTDCDSGKVDG---GDGSG	491	

Db 507 ACQC-PEGHVLRSKGKTC--AKLDSALGDHG 535

RESULT 13
NTC4_MOUSE

ID NTC4_MOUSE STANDARD; PRT; 1964 AA.

AC P31695; Q62389;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1997 (Rel. 35, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING PROTEIN INT-3).

GN NOTCH4 OR INT3 OR INT-3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92194507; PubMed=1312643;

RA Robbins J., Blondel B.J., Gallahan D., Callahan R.;

RT "Mouse mammary tumor gene int-3: a member of the notch gene family transforms mammary epithelial cells.";

RL J. Virol. 66:2594-2599(1992).

RN [2]

RP REVISIONS, SEQUENCE FROM N.A.

RX MEDLINE=97294599; PubMed=9150355;

RA Callahan D., Callahan R.;

RT "The mouse mammary tumor associated gene INT3 is a unique member of the NOTCH gene family (NOTCH4).";

RL Oncogene 14:1883-1890(1997).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung, and Testis;

RX MEDLINE=96281668; PubMed=8681805;

RA Uyttendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;

RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial cell-specific mammalian Notch gene.";

RL Development 122:2251-2259(1996)

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.

CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.

CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M80456; AAB38377.1; -.

DR EMBL; U43691; AAC52630.1; -.

DR PIR; A38072; TWVW13.

DR HSSP; P00740; IIXA.

DR MGD; MGI:107471; Notch4.

DR InterPro; IPR002110; ANK.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001438; EGF_II.

DR InterPro; IPR000800; Notch.

DR Pfam; PF00023; ank; 6.

DR Pfam; PF00008; EGF; 27.

DR Pfam; PF00066; notch; 2.

DR PRINTS; PR00010; EGFBLD.

DR SMART; SM00248; ANK; 5.

DR SMART; SM00179; EGF_CA; 11.

DR SMART; SM00001; EGF_like; 15.

DR SMART; SM00004; NL; 2.

DR PROSITE; PS50088; ANK_REPEAT; 5.

DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 11.

DR PROSITE; PS00022; EGF_1; 28.

DR PROSITE; PS01186; EGF_2; 21.

DR PROSITE; PS01187; EGF_CA; 9.

KW Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane; Glycoprotein; Proto-oncogene; ANK repeat; Signal.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 1964 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4.

FT DOMAIN 21 1443 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1444 1464 POTENTIAL.

FT DOMAIN 1465 1964 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 21 60 EGF-LIKE 1.

FT DOMAIN 61 112 EGF-LIKE 2.

FT DOMAIN 115 152 EGF-LIKE 3.

FT DOMAIN 153 189 EGF-LIKE 4.

FT DOMAIN 191 229 EGF-LIKE 5.

FT DOMAIN 231 271 EGF-LIKE 6.

FT DOMAIN 273 309 EGF-LIKE 7.

FT DOMAIN 311 350 EGF-LIKE 8.

FT DOMAIN 352 388 EGF-LIKE 9.

FT DOMAIN 389 427 EGF-LIKE 10.

FT DOMAIN 429 470 EGF-LIKE 11.

FT DOMAIN 472 508 EGF-LIKE 12.

FT DOMAIN 510 546 EGF-LIKE 13.

FT DOMAIN 548 584 EGF-LIKE 14.

FT DOMAIN 586 622 EGF-LIKE 15.

FT DOMAIN 622 656 EGF-LIKE 16.

FT DOMAIN 658 686 EGF-LIKE 17.

FT DOMAIN 688 724 EGF-LIKE 18.

FT DOMAIN 726 762 EGF-LIKE 19.

FT DOMAIN 764 800 EGF-LIKE 20.

FT DOMAIN 803 839 EGF-LIKE 21.

FT DOMAIN 841 877 EGF-LIKE 22.

FT DOMAIN 878 924 EGF-LIKE 23.

FT DOMAIN 926 962 EGF-LIKE 24.

FT DOMAIN 964 1000 EGF-LIKE 25.

FT DOMAIN 1002 1040 EGF-LIKE 26.

FT DOMAIN 1042 1081 EGF-LIKE 27.

FT DOMAIN 1083 1122 EGF-LIKE 28.

FT DOMAIN 1126 1167 EGF-LIKE 29.

FT REPEAT 1168 1208 LIN/NOTCH 1.

FT REPEAT 1209 1242 LIN/NOTCH 2.

FT REPEAT 1243 1282 LIN/NOTCH 3.

FT REPEAT 1628 1657 ANK 1.

FT REPEAT 1661 1691 ANK 2.

FT REPEAT 1695 1724 ANK 3.

FT REPEAT 1728 1757 ANK 4.

FT REPEAT 1761 1790 ANK 5.

FT DISULFID 25 38 BY SIMILARITY.

FT DISULFID 32 48 BY SIMILARITY.

FT DISULFID 50 59 BY SIMILARITY.

FT DISULFID 65 77 BY SIMILARITY.

FT DISULFID 71 100 BY SIMILARITY.

FT DISULFID 102 111 BY SIMILARITY.

FT DISULFID 119 130 BY SIMILARITY.

FT DISULFID 124 140 BY SIMILARITY.

FT DISULFID 142 151 BY SIMILARITY.

FT DISULFID 157 168 BY SIMILARITY.

FT DISULFID 162 177 BY SIMILARITY.

FT DISULFID 179 188 BY SIMILARITY.

FT DISULFID 195 208 BY SIMILARITY.

FT DISULFID 202 217 BY SIMILARITY.

FT DISULFID 219 228 BY SIMILARITY.

FT DISULFID 235 246 BY SIMILARITY.

FT DISULFID 240 259 BY SIMILARITY.

FT DISULFID 261 270 BY SIMILARITY.

FT DISULFID 277 288 BY SIMILARITY.

FT DISULFID 282 297 BY SIMILARITY.

FT DISULFID 299 308 BY SIMILARITY.

FT DISULFID 315 329 BY SIMILARITY.

FT DISULFID 323 338 BY SIMILARITY.

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
CC	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BALB/C; TISSUE-Limb;
RA	MEDLINE=97238863; PubMed=9083061;
RR	Deak F., Piccha D., Bachrati C., Karlsson M., Kiss I.;
RT	"Primary structure and expression of matrilin-2, the closest relative
RT	of cartilage matrix protein within the von Willebrand factor type A-
RT	like module superfamily.";
RL	J. Biol. Chem. 272:9268-9274(1997).
CC	-1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: SECRETED.
CC	-1- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
CC	CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
CC	OSTEOBLAST CELL LINES.
CC	-1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
CC	-1- SIMILARITY: CONTAINS 2 VWFA DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
EMBL	U69262; AAC53163.1; -.
DR	HSSP; P07204; 1ZAQ.
DR	MGI; 109613; Matn2.
DR	InterPro; IPR000152; Asx_hydroxyl.
DR	InterPro; IPR000561; EGF-like.
DR	InterPro; IPR001881; EGF-Ca.
DR	InterPro; IPR002035; VWFA.
DR	Pfam; PF00008; EGF; 10.
DR	Pfam; PF00092; vwa; 2.
DR	PRINTS; PR00453; VWFADOMAIN.
DR	SMART; SM00181; EGF; 10.
DR	SMART; SM00327; VWA; 2.
DR	PROSITE; PS00010; ASX_HYDROXYL; 7.
DR	PROSITE; PS01186; EGF_2; 9.
DR	PROSITE; PS0234; VWFA; 2.
KW	EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil.
FT	SIGNAL 1 23 POTENTIAL;
FT	CHAIN 24 956 MATRILIN-2.
FT	DOMAIN 57 232 VWFA 1.
FT	DOMAIN 238 278 EGF-LIKE 1.
FT	DOMAIN 279 319 EGF-LIKE 2.
FT	DOMAIN 320 360 EGF-LIKE 3.
FT	DOMAIN 361 401 EGF-LIKE 4.
FT	DOMAIN 402 442 EGF-LIKE 5.
FT	DOMAIN 443 483 EGF-LIKE 6.
FT	DOMAIN 484 524 EGF-LIKE 7.
FT	DOMAIN 525 565 EGF-LIKE 8.
FT	DOMAIN 566 606 EGF-LIKE 9.
FT	DOMAIN 607 647 EGF-LIKE 10.
FT	DOMAIN 655 830 VWFA 2.
FT	DOMAIN 917 955 COILED COIL (POTENTIAL).
FT	DISULFID 242 253 BY SIMILARITY.
FT	DISULFID 249 262 BY SIMILARITY.
FT	DISULFID 264 277 BY SIMILARITY.
FT	DISULFID 283 294 BY SIMILARITY.
FT	DISULFID 290 303 BY SIMILARITY.
FT	DISULFID 305 318 BY SIMILARITY.
FT	DISULFID 324 335 BY SIMILARITY.
FT	DISULFID 331 344 BY SIMILARITY.
FT	DISULFID 346 359 BY SIMILARITY.
FT	DISULFID 365 376 BY SIMILARITY.
FT	DISULFID 372 385 BY SIMILARITY.
FT	DISULFID 387 400 BY SIMILARITY.
FT	DISULFID 406 417 BY SIMILARITY.
FT	DISULFID 413 426 BY SIMILARITY.

